CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/62, C07K 14/54, 19/00, A61K 48/00, C12N 15/24, 15/83, A61K 38/20

A1

(11) International Publication Number: **WO 96/24676**

(43) International Publication Date: 15 August 1996 (15.08.96)

(21) International Application Number: PCT/US96/01787

(22) International Filing Date: 7 February 1996 (07.02.96)

(30) Priority Data:
08/385,335 8 February 1995 (08.02.95) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/385,335 (CIP)
Filed on 8 February 1995 (08.02.95)

(71) Applicant *(for all designated States except US)*: WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH [US/US]; Nine Cambridge Center, Cambridge, MA 02142 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LIESCHKE, Graham, J. [AU/US]; 5 Rollins Court, Cambridge, MA 02139 (US). MULLIGAN, Richard, C. [US/US]; 2 Sandy Pond Road, Lincoln, MA 01773 (US).

(74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AZ, BY, KG, KZ, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: BIOACTIVE FUSION PROTEINS AND PRE-EXISTING TUMOR THERAPY

(57) Abstract

Fusion proteins, such as a bioactive IL-12 polypeptide, which comprise at least two polypeptide monomers (chains of amino acids) joined through a heterologous polypeptide linker and which are bioactive, as well as their production, are described. Also described are tumor cells transduced to express the fusion proteins and methods of treating disorders characterized by established tumors.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

BIOACTIVE FUSION PROTEINS AND PRE-EXISTING TUMOR THERAPY

Related Applications

This application is a Continuation-in-Part of U.S. Serial No. 08/385,335, filed February 8, 1995, the teachings of which are incorporated herein by reference.

Background of the Invention

Production of therapeutic proteins, such as those which are dimeric, is often difficult, inefficient and expensive. Production of a dimer may require separate expression of the two components, followed by joining of those components to form a functional dimer. Alternative methods of producing functional dimeric proteins would be useful.

Summary of the Invention

The present invention relates to fusion proteins which comprise at least two polypeptide monomers (chains of amino acids) joined through a polypeptide linker and are bioactive, as well as to their production. In one embodiment, the bioactive fusion proteins of the present invention comprise two or more polypeptides which occur as subunits or monomers in a corresponding bioactive native dimeric protein and are linked through heterologous amino acid residues (amino acid residues which are not present between two subunits in the native protein). As it occurs in nature, the cytokine IL-12 is a heterodimer made up of a 40 kDa subunit (p40) linked by a disulfide bond to a 35 kDa subunit (p35). Gillessen. S. et al., Eur. J. Immunology, 25:200-206 (1995); Ozmen et al., J. Exp. Med., 180:907-915 (1995); Heinsel et al., Inf. & Immun., 62(10):4244-4249

(1994). For example, the fusion protein is a bioactive interleukin-12 (IL-12) fusion protein which comprises two subunits, designated p35 and p40, joined by a polypeptide linker. In further embodiments, the fusion protein comprises the subunits of other dimeric hematopoietic growth factors joined by a polypeptide linker, or the subunits of other dimeric cytokine proteins joined by a polypeptide linker. In another embodiment, the bioactive fusion protein comprises two subunits which are bioactive monomers (e.g., interleukin-2, GMCSF) in their native form and are joined through a polypeptide linker to produce a fusion protein which is chimeric or hybrid in nature in that it comprises at least two components or subunits which do not occur together in a native protein (e.g., an interleukin-2/GMCSF fusion protein).

The present invention also relates to methods of producing the subject fusion proteins, constructs useful in their production and host cells containing the constructs from which the encoded fusion proteins are expressed. The subject fusion proteins are expressed in an appropriate expression system, such as by a retrovirus vector which contains and expresses DNA encoding the subunits or monomers and the polypeptide linker of the desired fusion protein in an appropriate host cell, such as in mammalian cells. The present invention further relates to cells which have been transduced to secrete IL-12 fusion proteins of the present invention, and particularly to tumor cells which have been transduced to secrete an IL-12 fusion protein. The invention also relates to the use of the transduced tumor cells, particularly in the treatment of tumors.

Fusion proteins of the present invention are useful for the same purposes (e.g., therapeutic or diagnostic uses) as the corresponding native protein. For example, IL-12 fusion proteins can be used to enhance the lytic

activity of NK/lymphokine - activated killer cells, act as a growth factor for activated human T and NK cells and stimulate production of IFN-γ by resting peripheral blood mononuclear cells (PBMC). IL-12 is also useful in treating a variety of cancers. For instance, IL-12 is useful for the enhancement of antitumor immunity and, as described herein, tumor cells which secrete either native IL-12 or an IL-12 fusion protein of the present invention can be used to treat established tumors, such as to prevent the further development of a tumor, cause established tumors to regress, prolong survival, or a combination thereof. The fusion proteins have certain advantages over the corresponding native proteins in that they can be made efficiently and reproducibly by the methods described herein. Furthermore, the fusion proteins of the present invention may also have advantages over the corresponding native proteins in terms of modified or enhanced activity, more favorable bioavailability and improved pharmacokinetic properties.

Brief Description of the Drawings

Figure 1 shows the structures of SFG-based retroviral constructs for interleukin-12 production (SD=splice donor; IRES=internal ribosome entry site; SA=splice acceptor; LTR=long terminal repeat).

Figure 2 shows the nucleic acid sequences encoding the linker sequences in interleukin-12 fusion proteins of the present invention and flanking IL-12 p35 and Il-12 p40 sequences (SEQ ID NO: 1 to 4 and 35), as well as the encoded amino acid sequences (SEQ ID NO: 5 to 7 and 36).

Figures 3A-3U show the full restriction map and the nucleic acid sequence (SEQ ID NO: 8 and 9) of pUC19-SFG.

Figures 4A-4C show the nucleic acid sequence (SEQ ID NO: 10 and 11) encoding the murine IL-12 p35 subunit and

the amino acid sequence of the murine IL-12 p35 subunit (SEQ ID NO: 12).

Figures 5A-5D show the nucleic acid sequence (SEQ ID NO: 13 and 14) encoding murine IL-12 p40 subunit and the amino acid sequence (SEQ ID NO: 15) of the murine IL-12 p40 subunit.

Figure 6 shows a standard curve generated using recombinant murine IL-12.

Figures 7A-7D show graphic representations of the effect of immunotherapy of CMS-5 tumor-bearing mice with wild-type, GM-CSF- and IL-12-secreting CMS-5 cells. Treatment was started either on day 7 (7A and 7B) or day 14 (7C and 7D) after tumor challenge. Endpoints are either survival (7A and 7C) or tumor-free survival (7B and 7D). Tumors were either untreated (a) or treated with GM-CSF-secreting CMS-5 cells (b), IL-12-secreting CMS-5 cells (c) or wild type CMS-5 cells (d).

Figure 8 is a graphic representation of the incidence of regression of established CMS-5 tumors by type of immunotherapy. Tumors were treated as follows: column 1 received no immunotherapy; column 2 was treated with wild-type tumor cells; column 3 was treated with GM-CSF-secreting tumor cells; and column 4 was treated with IL-12-secreting tumor cells.

Figure 9 is a graphic representation of tumor regression in mice treated with systemic IL-12. The open square and closed square, triangle, circle and diamond represent 5 individual mice treated with systemic IL-12 at 0.1 μg/d given 5 days per week for 4 weeks.

Figures 10A-10B show graphic representations of the superior survival resulting from immunotherapy with IL-12-secreting CMS-5 cells compared to systemic IL-12 administration or no treatment (nil). Figure 10A depicts results using a tumor inoculum of 2×10^5 cells. Figure

10B depicts results using a tumor inoculum of 4 x 10^5 cells.

Figures 11A-11B show graphic representations of the comparison of efficacy (proportion of mice surviving) of CMS-5 cells secreting different forms of IL-12 as immunotherapy for established CMS-5 tumors. Figure 11A shows results using tumors initiated by 2 x 10^5 CMS-5 cells with treatment starting on day 14 (20 mice per group pooled from two experiments). Figure 11B shows results using tumors initiated by 4 x 10^5 CMS-5 cells with treatment starting on day 14 (1 group of 10 mice). The tumors were either untreated (a) or treated with wild type CMS-5 cells (b), GM-CSF-secreting CMS-5 cells (c), native IL-12-secreting CMS-5 cells (d) or IL-12 fusion protein-secreting CMS-5 cells (e).

Figures 12A-12C are graphic illustrations of the results of immunotherapy of B16 (melanoma) tumors with cytokine-secreting tumor cells. For the pre-existing tumor model, tumors were initiated with 4 x 10^5 B16 cells and immunotherapy commenced on day 7 (Figure 12A) or day 14 (Figure 12B). For the challenge model (Figure 12C), 5 x 10^5 irradiated cells were administered as a vaccine 14 days before tumor challenge with 1 x 10^6 B16 cells. Tumors were either untreated (a) or treated with wild type B16 cells (b), GM-CSF-secreting B16 cells (c), native IL-12-secreting B16 cells (d) or IL-12 fusion protein-secreting B16 cells (e).

Figures 13A-13B are graphic illustrations of the effects on immunotherapy of IL-12 delivery by different cell types in mice with pre-existing renal cell carcinoma (RENCA) tumors. Figure 13A shows results when RENCA tumors were treated with either irradiated wild-type CMS-5 tumor cells (C-wt) or CMS-5 tumor cells transduced to secret either native IL-12 (C-nIL-12) or the IL-12 fusion protein (C-scIL-12). Figure 13B shows results when RENCA tumors

were treated with either a combination of wild-type CMS-5 and RENCA cells (C-wt + R-wt), a combination of IL-12 fusion protein-secreting RENCA cells and wild type CMS-5 cells (C-wt + R-IL-12) or a combination of IL-12 fusion protein-secreting CMS-5 cells and wild type RENCA cells (C-IL-12 + R-wt).

Figures 14A-14B are a graphic representation of the effects on immunotherapy of pre-existing CMS-5 tumors with IL-12 fusion protein-secreting RENCA tumor cells. Figure 14A shows the results when CMS-5 tumors were treated with either wild type RENCA cells or RENCA cells transduced to secrete the IL-12 fusion protein. Figure 14B shows the results when CMS-5 tumors were treated with either a combination of wild type RENCA and wild type CMS-5 cells (C-wt + R-wt), a combination of IL-12 fusion protein-secreting RENCA cells and wild type CMS-5 cells (C-wt + R-IL-12) or a combination of IL-12 fusion protein-secreting CMS-5 cells and wild type RENCA cells (C-IL-12 + R-wt).

Detailed Description of the Invention

Described herein are bioactive fusion proteins which comprise at least two subunits linked or joined by an intervening amino acid linker, a method of producing the bioactive fusion proteins, constructs useful for producing the fusion proteins which can be expressed in host cells, and host cells containing the constructs.

In one embodiment, the bioactive fusion proteins of the present invention comprise: 1) at least two polypeptide subunits or monomers which correspond to polypeptide subunits present in a native dimeric protein which has a specified bioactivity and 2) at least one polypeptide linker which joins the subunits in such a manner that the resulting fusion protein is bioactive. If the resulting fusion protein is dimeric (includes two subunits or monomers), the two components can be subunits which occur

in the same native dimeric protein (e.g., two IL-12 subunits); subunits which occur in two different native dimeric proteins (e.g., one subunit from IL-12 and one subunit from IL-3) or monomers which are individually bioactive (e.g., IL-2, GMCSF). Multimeric fusion proteins, which comprise three or more subunits joined by polypeptide linkers, can comprise, for example, three or more of the subunits which occur in the same native dimeric protein (e.g., three or more IL-12 subunits), three or more subunits which occur in different native dimeric proteins (e.g., two IL-12 subunits and one IL-3 subunit), three or more bioactive monomers (e.g., three IL-2 monomers, two IL-2 monomers and one GMCSF monomer) or a combination of subunits from native dimeric proteins and bioactive monomers (e.g., two IL-12 subunits and a GMCSF monomer). In each case, a polypeptide linker is present between two subunits (e.g., the order is subunit-linker-subunit-linker-subunit). As used herein, the terms subunit and monomer are used interchangeably to refer to the components of a dimeric or multimeric protein and the single component of a monomeric protein. The order of subunits in the fusion protein of the present invention can be p35-linker-p40 or p40-linker-p35. In either case, the polypeptide linker is positioned between the two subunits. A bioactive fusion protein of the present invention which includes subunits which occur in the same native dimeric protein "mimics" or is similar to what is referred to herein as a corresponding native dimeric protein in terms of its bioactivity, but differs from the corresponding native dimeric protein in that the fusion protein includes linker amino acid residues which do not occur in the corresponding native protein (heterologous amino acid residues) between each pair of polypeptide subunits. A corresponding native protein is one which includes the subunits present in the fusion

protein and exhibits biological activity also exhibited by the fusion protein.

For example, in the case of a bioactive IL-12 fusion protein, the two subunits, designated p35 and p40, of a mammalian native IL-12 protein (e.g., human, mouse, rat, dog, cat, monkey, chimpanzee or pig IL-12 protein) are joined through a polypeptide linker. Here, the corresponding native protein is the mammalian native IL-12 protein. Similarly, in the case of another bioactive fusion protein, such as IL-3, the corresponding native protein is IL-3. The amino acid residues of the subunits of the bioactive fusion protein can be the same as those of the subunits of the corresponding native protein or can be different, provided that the resulting fusion protein exhibits the desired bioactivity. For example, the subunit(s) can have a different amino acid sequence from that of the corresponding subunit of a native protein (i.e., the sequence of the native subunit can differ in that one or more amino acid residues has been deleted or replaced by a naturally-occurring or non-naturally-occurring amino acid residue, additional amino acid residues have been incorporated, or an amino acid residue has been modified). The desired bioactivity is activity like that of the corresponding native protein (e.g., it produces a physiological response which also results from the activity of the corresponding native protein). The bioactivity of a fusion protein (e.g., the duration of its effect, extent of the resulting response) may be greater or lesser than that of the corresponding native protein.

The polypeptide linker present in the fusion protein can be of any length and composition appropriate to join two subunits in such a manner that the resulting fusion protein has the desired biological activity and retains its integrity as a dimer or multimer. The appropriate length and composition of a linker can be determined empirically

for the specific fusion protein to be produced. Generally, the polypeptide linker will be at least 7 amino acid residues, although it can be shorter (e.g., 2 to 6 amino acid residues). Typically the linker will be less than 30 amino acid residues in length, such as 7 to 25 amino acid residues or 7 to 20 amino acid residues in length. In one embodiment, the polypeptide linker is 7 to 16 amino acid residues and in specific embodiments is 7, 11, 15 or 16 amino acid residues. Specific linkers used in producing bioactive IL-12 fusion proteins are represented in Figure 2 and described in Example 4. In specific embodiments, the polypeptide linkers are exemplified by the sequences $(Gly_4Ser)_3$; $(Gly_4Ser)_3Ser$; $(Gly_4Ser)_2Ser$ and Gly_6Ser, and these linkers can also be used to join subunits of other fusion proteins in addition to the IL-12 fusion proteins of the present invention. Alternatively, other polypeptide linkers can be used to join two IL-12 subunits to produce a bioactive IL-12 fusion protein.

The DNA encoding the bioactive fusion protein can be cDNA or genomic DNA and can be from a variety of animals, particularly mammals. For example, the DNA can be human, mouse, rat, dog, cat, monkey, chimpanzee, pig or ferret DNA. The DNA can encode a complete or entire subunit (e.g., a complete IL-12 p35 subunit and a complete IL-12 p40 subunit) or a fragment or portion of a subunit(s), provided that the encoded fusion protein has the desired biological activity when it is expressed. The nucleic acid sequences of DNA encoding mouse IL-12 p35 and p40 subunits are represented in Figures 4 and 5, respectively. The nucleic acid sequences of DNA encoding human IL-12 p35 and p40 subunits have been published (Gubler et al. in Proceedings of the National Academy of Sciences, USA, 88:4143 (1991); Figure 4A-4C and 5A-5D). All or a portion of IL-12 DNA can be used to produce the subject IL-12

fusion protein, provided that the encoded fusion protein is bioactive (has IL-12 activity).

Any expression system appropriate for expressing a protein such as a mammalian, bacterial, yeast or insect expression system, can be used to express the fusion proteins of the present invention. For example, as described herein, a viral (e.g., a retroviral) vector which expresses DNA (e.g., cDNA) encoding the desired fusion protein in a mammalian host cell has been used. As also described herein, retroviruses containing cDNA encoding the p35 and p40 subunits of IL-12 and an intervening polypeptide linker (an IL-12 fusion protein) have been constructed and transfected into packaging cells (e.g., BOSC23 packaging cells). Target cells (e.g., CMS-5 fibrosarcoma cell line) were infected with virus-containing supernatants and cultured; media conditioned by infected cells was assayed for IL-12 activity using an interleukin-2 and concanavalin-A primed splenocyte proliferation bioassay. Packaging or producer cell lines other than BOSC23 cells can be used to produce infectious retroviruses containing the fusion protein-encoding DNA. In addition, target cells other than a fibrosarcoma cell line, such as B16 melanoma or renal cell carcinoma cell lines, can be used to produce the fusion protein. IL-12 bioactivity was demonstrable in cells infected with the retroviruses, as described in Example 4.

Specific retroviruses have been constructed for expression of an IL-12 fusion protein (Example 1 and Figure 1) and cells infected with the retroviruses have been shown to produce bioactive IL-12 fusion proteins (see Example 4). The retroviruses used all included the SFG retroviral backbone whose sequence is shown in Figure 3. The vectors designated pSFG.IL-12.p35 and pSFG.IL-12.p40 include, respectively, the cDNA for the IL-12 p35 subunit or the cDNA for the IL-12 p40 subunit. The vector designated

pSFG.IL-12p35-IRES-p40 includes cDNA encoding the IL-12 p35 subunit and cDNA encoding the IL-12 p40 subunit, separated by an internal ribosome entry site sequence. The vector designated pSFG.IL-12p40-IRES-p35 includes the same components as plasmid pSFG.IL-12p35-IRES-p40 but the dimers are in the reverse order, as indicated. The vectors designated pSFG.IL-12.p35-linker-p40 and pSFG.IL-12.p40-linker-p35 include cDNAS encoding each IL-12 subunit linked by the $(Gly_4Ser)_2Ser$ and $(Gly_4Ser)_3Ser$ linker respectively. The vectors designated pSFG.IL-12.p35-linker-Δp40 and pSFG.IL-12.p40-linker-Δp35 include linked cDNAs in which sequences encoding a putative 22 amino acid leader sequence were deleted from the second cDNA. The vector designated pSFG.hIL-12.p40.linker.Δp35 is a human form of the IL-12 fusion protein and is analogous to the murine form pSFG.IL-12.p40.linker.Δp35 except that the linker is shorter due to a deletion which occurred during the cloning (see Figure 2, construct E). As described in Example 4, IL-12 bioactivity was shown in conditioned medium from cells infected with the retroviruses.

Prokaryotic and eukaryotic host cells transfected by the described vectors are also provided by this invention. For instance, cells which can be transfected with the vectors of the present invention include, but are not limited to , bacterial cells such as *E. coli*, insect cells (baculovirus), yeast or mammalian cells such as Chinese hamster ovary cells (CHO). Tumor cells which are transduced to secrete the IL-12 fusion proteins of the present invention are particularly useful in the present invention.

Thus, expression vectors described herein can be used to transform, transfect or transduce host cells, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial cells), using standard procedures

used in producing other well known proteins. Similar procedures, or modifications thereof, can be employed to prepare recombinant proteins according to the present invention by microbial means or tissue-culture technology. For example, fibroblast-derived 3T3 cells can be transduced with the vectors of the present invention to express and secrete the IL-12 fusion proteins of the present invention. Tumor cells and 3T3 cells are useful in the context of the present invention, as cells transduced to secrete IL-12 or IL-12 fusion proteins of the present invention are a useful source of the protein or fusion protein (e.g., for purification). Tumor cells transduced to secrete IL-12 or IL-12 fusion proteins also have particular utility as antitumor agents as described herein.

The tumor cells to be transduced can be selected from the individual to be treated or from another individual; furthermore, the tumor cells to be transduced can be the same type as the tumor cells of the tumor to be treated or the tumor cells can be of a different type from the tumor to be treated. For example, a CMS-5 tumor can be treated with CMS-5 tumor cells, renal cell carcinoma (RENCA) tumor cells, or B16 tumor cells which secrete IL-12 or an IL-12 fusion protein of the present invention. Alternatively, the tumor can be treated with a combination of IL-12-secreting, IL-12 fusion protein-secreting and wild type cells of the same or different cell type. For example, a RENCA tumor can be treated with a combination of wild type RENCA cells and IL-12 fusion protein-secreting CMS-5 cells or with a combination of native IL-12-secreting CMS-5 cells and IL-12 fusion protein-secreting RENCA tumor cells.

The present invention also relates to transduced tumor cells which express native IL-12 or IL-12 fusion proteins of the present invention and their use in treating tumors. That is, transduced tumor cells which express and secrete IL-12 or IL-12 fusion proteins are useful as therapeutic

agents for the treatment of cancer or to treat established tumors, and provide a means for reversing tumors (reducing their size or causing their complete regression) or preventing further growth of an established tumor. As described herein, transduced tumor cells expressing IL-12 or IL-12 fusion proteins of the present invention cause the regression of established tumors, prevent the establishment of tumors, prolong survival or a combination thereof in animals to which they are administered in a therapeutically appropriate dose.

For instance, the tumor cells secreting IL-12 or the IL-12 fusion protein of the present invention can be formulated with a physiologically acceptable medium to prepare a pharmaceutical composition. The particular physiological medium may include, but is not limited to, water, buffered saline, polyols (e.g., glycerol, propylene glycol, liquid polyethylene glycol) and dextrose solutions. The optimum concentration of the active ingredient(s) in the chosen medium can be determined empirically, according to procedures well known to medicinal chemists, and will depend on the ultimate pharmaceutical formulation desired. Methods of introduction of the IL-12-secreting or IL-12 fusion protein-secreting tumor cells include, but are not limited to, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, oral and intranasal. The tumor cells secreting native IL-12 or an IL-12 fusion protein can be administered at or near the site of a tumor to be treated or at any other site on the body, provided that the IL-12 or IL-12 fusion protein produces the desired therapeutic effect (regression of established tumors, prevention of tumor establishment or prolonged survival). As described herein, proximity of tumor site and administration site is not necessary for the efficacy of the treatment. Other suitable methods of introduction can also include rechargeable or biodegradable devices and slow

release polymeric devices. The pharmaceutical compositions of this invention can also be administered as part of a combinatorial therapy with other agents. Treatment regimens will depend upon the dose, route of delivery, frequency with which the composition is administered, type, size and stage of the tumor to be treated, and the age, health and other physical characteristics of the individual to be treated.

In one method of the invention, a therapeutically effective quantity (dose) of the transduced tumor cells (optionally formulated with a physiologically appropriate medium) is administered to an individual having a tumor to be treated (e.g., decreased in size or prevented from increasing in size). The transduced tumors cells can also be administered in a therapeutically appropriate dose to an individual to prevent the establishment of a tumor. For example, the transduced tumor cells can be administered to an individual in need of anticancer therapy (e.g., an individual with an established tumor or an individual in whom establishment of a tumor is to be prevented). Alternatively, the IL-12 fusion protein of the present invention can be administered directly to the individual in a therapeutically effective dose; the IL-12 fusion protein can be optionally combined with a physiologically acceptable medium as described above.

As described herein, the efficacy of IL-12-secreting tumor cells as antitumor immunotherapy was assessed in mice with established tumor burdens. The immunogenic CMS-5 (fibrosarcoma) and non-immunogenic B16 (melanoma) tumors were used; RENCA tumors were also utilized as described herein.

As shown in Examples 6-8, work described herein demonstrates that for the immunotherapeutic treatment of 14-day established palpable CMS-5 tumors, immunotherapy with IL-12-secreting and IL-12 fusion protein-secreting

tumor cells prolonged survival by inducing the regression of tumors. Furthermore, immunotherapy with IL-12-secreting tumor cells induced tumor regression even when the palpable tumor burden averaged more than 5% of body mass. Although IL-12 has antitumor activity against CMS-5 tumors when administered systemically, for mice with larger tumor burdens at the onset of therapy there was a significant survival advantage of IL-12-secreting tumor cell immunotherapy over systemic IL-12 therapy. This shows that there is an advantage in delivering the IL-12 by transduced tumor cells rather than just as systemic cytokine therapy. Data from tumor cells transduced to express an IL-12 fusion protein (SFG.IL-12.p40.linker.Δp35) indicate that the murine and human forms of the fusion protein have a specific activity at least equal to the native molecule in an in vitro bioassay.

Results described herein show that IL-12-secreting B16 cell vaccination altered the natural history of the growth of later established B16 tumors, and they also appear to be able to enhance immunological mechanisms capable of modulating tumor growth. IL-12-secreting B16 cells are useful as immunotherapy for established B16 tumors, as they effectively prolong survival. These results show that there exist inducible innate mechanisms able to modulate the natural history of established tumors in the mouse, and that IL-12-secreting cells are more potent at inducing them than GM-CSF-secreting cells. These results, which are more fully described in the Examples below, show that IL-12-secreting tumor cells have efficacy as immunotherapy for established tumors.

The present invention is illustrated by the following examples, which are not intended to be limiting in any way.

EXAMPLE 1 Construction of Plasmids

The general structure of the plasmids used in these studies is shown schematically in Figure 1. The confirmed sequences of the linkers in each of the fusion proteins are given in Figure 2.

Source of plasmids

The plasmids containing cDNAs for the murine IL-12 p35 and p40 subunits (pBS.IL-12.p35 and pBS.IL-12.p40) were provided by Hoffmann-La Roche (Nutley, NJ). The numbering of base pairs in this document corresponds to the maps of the inserts of these two plasmids (Figures 4 and 5). The plasmid containing the SFG retroviral backbone was provided by Dr. Dan Ory (Whitehead Institute, Cambridge, MA) as pSFG-TPA, a pUC plasmid containing the SFG retroviral backbone between the HindIII and EcoR1 sites with a tissue plasminogen activator cDNA between the unique Nco1 and BamH1 sites in the SFG retrovirus. A nucleotide sequence map of the SFG retroviral backbone is shown in Figure 3.

Plasmid pSFG.IL-12.p35

The IL-12p35 cDNA was provided in pBluescript with the sequences surrounding the translational initiation ATG optimized to ACCATGG according to the rules of Kozak. The IL-12p35 cDNA fragment was excised as a Nco1-EcoR1 fragment, the EcoR1 overhang having been filled using the Klenow fragment of *E. coli* DNA polymerase 1. This fragment was ligated using T4 DNA ligase into the Nco1-BamH1 sites of pSFG, the BamH1 overhang having been filled using the Klenow fragment of *E. coli* DNA polymerase 1. The resulting plasmid is designated pSFG.IL-12.p35.

Plasmid pSFG.IL-12.p40

The IL-12p40 cDNA was provided in pBluescript. The Nco1-BamH1 fragment containing the IL-12p40 cDNA was

excised and ligated into the Nco1-BamH1 sites of pSFG to make pSFG.IL-12.p40.

General Strategy for Construction of SFG-based Vectors

The general strategy for constructing the SFG-based retroviral vectors for IL-12 fusion protein expression is as follows: Two oligonucleotides encoding the sense and antisense strand of a $(Gly_4Ser)_3$ linker fragment and contiguous IL-12 cDNA sequences to be linked (with terminal sequences for the creation of cohesive ligatable overhangs) were synthesized using a "PCR-mate" 391 DNA synthesizer (Applied Biosystems, Foster City, CA). The sequence of the $(Gly_4Ser)_3$ linker was that of Huston et al. (Proc. Natl. Acad. Sci. USA, 85:5879-5883(1988)).

For the two fusion proteins using complete IL-12 cDNAs, the oligonucleotides were designed to be cloned into a unique restriction enzyme site at the 3' end of the first cDNA, reconstructing the 3' end of the first cDNA and enabling a Nco1-Nco1 fragment encompassing the full cDNA and linker sequence to be cloned into the Nco1 site of the SFG plasmid containing the other cDNA.

The cloning strategy was similar for the two fusion proteins with a deletion of 66 bp coding the first 22 amino acids of the second cDNA. Linker oligonucleotides were designed to be cloned into unique restriction enzyme sites that lay 3' of bp 66 of the translated bases of the second cDNA in the fusion protein construct. This enabled a fragment to be excised for cloning that reconstructed the 3' end of the first cDNA joined to the linker and contained the linker joined to codon 23 of the second cDNA.

The sequence of the linker and contiguous cDNA regions in plasmids was determined using a "Sequenase" kit (Amersham, Cleveland, OH).

Plasmid pSFG.IL-12.p35-linker-p40

The oligonucleotides were: sense, 5'-CCGCC.GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.TCCATGGAGCT-3' (SEQ ID NO: 16); and antisense, 5'-CCATGGA.AGA.TCC.GCC.GCC.ACC.CGA.CCC.ACC.ACC.GCC.CGA.GCC.ACC.GCC.ACC.GGCGGAGCT-3' (SEQ ID NO: 17).

These two oligonucleotides were annealed, phosphorylated using T4 polynucleotide kinase, and ligated into the Sac1 site of pBS.IL-12.p35 which had been dephosphorylated using calf intestinal phosphatase. The Nco1-Nco1 fragment of the resulting plasmid containing the IL-12p35 cDNA and correctly orientated linker was excised and ligated into the dephosphorylated Nco1 site of pSFG.IL-12p40 to create pSFG.IL-12.p35-linker.p40 (the correct orientation of this ligated fragment was demonstrated by a Sac1 digest).

This plasmid was sequenced using the following two primers: 5'-CAGAGTGAAAATGAAGCT-3' (SEQ ID NO: 18) and 5'-GAAGCTCTGCATCCTGCT-3' (SEQ ID NO: 19), corresponding to bp 601-618 and 613-630 of the IL-12p35 cDNA. Sequencing demonstrated that a deletion had occurred during cloning resulting in a loss of 15 bp from the linker sequences, but maintaining an intact reading frame. The sequence of the linker in this plasmid is given in Figure 2.

Plasmid pSFG.IL-12.p40.linker.p35

The oligonucleotides were: sense, 5'-GGGTCCGATCC.GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.TCCATG-3' (SEQ ID NO: 20); and antisense, 5'-GATCCATGGA.AGA.TCC.GCC.GCC.ACC.CGA.CCC.ACC.ACC.GCC.CGA.GCC.ACC.GCC.ACC-GGATCGGACCCTGCA-3' (SEQ ID NO: 21).

These two oligonucleotides were annealed and ligated into the Sse83871 and BamH1 sites of pBS.IL-12.p40. The Nco1-Nco1 fragment of the resulting plasmid containing the IL-12p40 cDNA and correctly orientated linker was excised

and ligated into the dephosphorylated Nco1 site of pSFG.IL-12p35 to create pSFG.IL-12.p40.linker.p35 (the correct orientation of this ligated fragment was demonstrated by a Xcm1 digest).

This plasmid was sequenced using the following two primers: 5'-CTATTACAATTCCTCATG-3' (SEQ ID NO: 22) and 5'-GAGGGCAAGGGTGGCCAA-3' (SEQ ID NO: 23), corresponding to base pairs 997-1014 of the IL-12 p40 cDNA and base pairs 91-74 of the IL-12 p35 cDNA (an antisense primer). Sequencing confirmed that the sequence of the linker and contiguous IL-12 cDNA sequences were as expected.

Subsequent restriction enzyme mapping of pSFG.IL-12.p40.linker.p35 after the transfection and expression studies were completed revealed that it probably contained a concatamer of Nco1-Nco1-fragments from the final cloning step.

Plasmid pSFG.IL-12.p35.linker.Δp40

The oligonucleotides were: sense, 5'-T.TGC.TGG.AGC.TCC.GCC.GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GG G.TCG.GGT.GGC.GGC.GGA.TCT.ATG.TGG-3' (SEQ ID NO: 24) and antisense, 5'-CACAT.AGA.TCC.GCC.GCC.ACC.CGA.CCC.ACC.ACC.GCC.CGA.GCC.AC C.GCC.ACC.GGCGGAGCTCCAGCAAA-3' (SEQ ID NO: 25).

These two oligonucleotides were annealed, phosphorylated using T4 polynucleotide kinase, and ligated into pBS.IL-12.p40 from which the 30 base pair 5' Xcm1-Xcm1 fragment had been excised. The Sac1-Sac1 fragment from the resultant plasmid was excised and ligated into the Sac1 site of pBS.IL-12.p35 which had been dephosphorylated using calf intestinal phosphatase (the correct orientation of the ligated fragment was demonstrated by a Nco1-EcoR1 digest). The Nco1-EcoR1 fragment of the resultant vector was excised, the EcoR1 overhang having been filled using the Klenow fragment of *E. coli* DNA polymerase 1, and ligated

into the Nco1 and Klenow-filled BamH1 sites of pSFG to create pSFG.IL-12.p35.linker.Δp40.

This plasmid was sequenced using the following primers: 5'-CAGAGTGAAAATGAAGCT-3' (SEQ ID NO: 18) and 5'-GAAGCTCTGCATCCTGCT-3' (SEQ ID NO: 19), corresponding to base pairs 601-618 and 613-630 of the IL-12p35 cDNA; and 5'-GTCATCTTCTTCAGGCGT-3' (SEQ ID NO: 34), an antisense primer corresponding to base pairs 217-200 of the IL-12 p40 cDNA. Sequencing confirmed that the sequence of the linker and contiguous IL-12 cDNA sequences were as expected.

Plasmid pSFG.IL-12.p40.linker.Δp35

The oligonucleotides were: sense, 5'-CTG.GCC.TGC.AGG.GTC.CGA.TCC-GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT-AGG.GTC.ATT.CCA.GTC.T-3' (SEQ ID NO: 26) and antisense, 5'-CTGGAATGACCCT.AGA.TCC.GCC.GCC.ACC.CGA.CCC.ACC.ACC.GCC.CGA.GCC.ACC.GCC.ACC.GGATCGGACCCTGCAGGCCAGAGA-3' (SEQ ID NO: 27).

These two oligonucleotides were annealed, phosphorylated using T4 polynucleotide kinase, and ligated into the Pf1M1 site in pBS.IL-12.p35 which had been dephosphorylated using calf intestinal phosphatase. The orientation of this ligated fragment was confirmed by an Sse83871/EcoR1 digest. The Sse83871-EcoR1 fragment from the resultant plasmid was excised, the EcoR1 overhang having been filled using the Klenow fragment of *E. coli* DNA polymerase 1, and ligated into the Sse83871 and Klenow-filled BamH1 sites of pSFG.IL-12.p40 to create pSFG.IL-12.p40.linker.Δp35.

This plasmid was sequenced using the primer 5'-GCAAAGGCGGGAATGTCT-3' (SEQ ID NO: 28), corresponding to base pairs 960-977 of the IL-12.p40 cDNA. The sequence of the second linker codon was difficult to read, but its sequence was determined by sequencing the cloned linker in

the intermediate plasmid using the antisense primers 5'-AGGAATAATGTTTCAGTT-3' (SEQ ID NO: 29) and 5'-CAGCAGTGCAGGAATAAT-3' (SEQ ID NO: 30) corresponding to base pairs 224-207 and 233-216 of the IL-12 p35 cDNA respectively. Sequencing confirmed that the sequence of the linker and contiguous IL-12 cDNA sequences were as expected.

Plasmids pSFG.IL-12.p35.IRES.p40 and pSFG.IL-12.p40.IRES.p35

The encephalomyelocarditis virus (ECMV) internal ribosome entry site (IRES) fragment was provided by Dr. Michael Sadelain (Whitehead Institute, Cambridge, MA), and was as previously described (Ghattas et al., Mol. Cell. Biol., 11:5848-5859 (1991)).

EXAMPLE 2 Cells and Tissue Culture

BOSC23 packaging cells (Pear et al., Proc. Natl. Acad. Sci. USA, 90:8382-8396(1993)) were obtained from Dr. Dirk Lindemann (Whitehead Institute, Cambridge, MA). They were passaged in Dulbecco's modified Eagles medium (DMEM) supplemented with 10% calf serum, 50 U/ml penicillin and 50 μg/ml streptomycin.

CMS-5 tumor cells (DeLeo et al., J. Exp. Med., 146:720-734 (1977)) were obtained from Jason Salter (Whitehead Institute, Cambridge, MA). They were passaged in DMEM supplemented with 10% foetal calf serum, 50 U/ml penicillin and 50 μg/ml streptomycin. The same medium was used for the collection of CMS-5 conditioned medium.

C57BL/6 splenocytes for IL-12 assays were obtained by mincing a spleen through a sieve (Falcon 2350, Becton Dickinson, Franklin Lakes, NJ) and collecting the cells in IL-12 medium (as detailed in Schoenhaut et al. (J. Immunol., 148:3433-3440 (1992)) supplemented with 2% foetal calf serum.

EXAMPLE 3 Generation of BOSC23-derived Producer Cells and Collection of Conditioned Media

BOSC23 cells were plated at 2 x 10^6 cells per 6 cm tissue culture dish and transfected by $CaPO_4$ transfection with the various constructs as previously described (Pear et al., Proc. Natl. Acad. Sci. USA, 90:8382-8396 (1993)). Twenty-four hours after transfection, the medium was replaced with 5 ml fresh medium. Virus-containing supernatants were collected 24 hours later, filtered through a 0.45 μm filter and polybrene added to a final concentration of 8 μg/ml. 2.5 ml of virus-containing supernatant was used to infect CMS-5 cells immediately for 4 hours (in preparation for this infection, CMS-5 cells had been plated at 5x10^4 cells/6 cm tissue culture dish the previous day) and the remaining 2.5 ml frozen at -70 °C. The following day, the frozen 2.5 ml of virus-containing supernatant was thawed and used for a second 4 hour infection of the CMS-5 cells. To collect IL-12-containing conditioned medium, the medium was replaced the following day with 5 ml fresh medium which was harvested 24 hours later. These conditioned media were filtered through a 0.2 μm filter and frozen at -70°C for later assay for IL-12 bioactivity. 5 ml of fresh medium was added to the CMS-5 cells and a second set of conditioned media collected 24 hours later which were also filtered and frozen for later assay. The infected CMS-5 cells were then lysed, and genomic DNA prepared for later analysis.

EXAMPLE 4 Bioassay for Murine Interleukin-12

Levels of bioactive interleukin-12 were determined using a concanavalin-A and interleukin-2 primed splenocyte proliferation assay, as described in Schoenhaut et al. (J. Immunol., 148:3433-3440 (1992)). The concanavalin A was obtained commercially from Boehringer (Mannheim, Germany)

and the recombinant human interleukin-2 commercially from Chiron Therapeutics (Emeryville, CA). To harvest cells for the measurement of [^{3}H]thymidine incorporation into cellular DNA, a Skatron (Sterling, VA) cell harvester and filtermats (#7031) were used. To assay for inhibitory activity in conditioned media, the 50 μl sample volume comprised 25 μl of 1000 pg/ml recombinant murine IL-12 and 25 μl of the test sample. Samples of conditioned media were assayed in duplicate at several dilutions in the range 1:1 to 1:1000. A standard curve was constructed for each bioassay using recombinant murine IL-12 in the range 20-10,000 pg/ml. The recombinant murine IL-12 was obtained from Hoffmann-La Roche (Nutley, NJ). To calculate the bioactive IL-12 concentration in test samples in pg/ml, the linear part of the standard curve was approximated using the curve-fit function of "KaleidaGraph 2.1.1" software and the resultant formula used for calculations. Conditioned media were verified to have hIL-12 immunoreactivity by hIL-12 ELISA assay (commercial kit, R & D Systems).

The following constructs (Figure 1) were assessed for their ability to express a bioactive IL-12 fusion protein:

A. pSFG.IL-12.p35.linker.p40;
B. pSFG.IL-12.p40.linker.p35;
C. pSFG.IL12-p35.linker.Δp40;
D. pSFG.IL12-p40.linker.Δp35; and
E. pSFG.hIL-12.p40.linker.Δp35.

The sequences for the linkers in each construct were as follows, as confirmed by sequencing (some adjacent confirmed IL-12 sequences are given for orientation):

A. 5'->>>IL-12p35.AGC.TCC.GCC-GGT.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.TCC.ATG.GGT.CCT.CAG.>>>IL-12p40-3' (SEQ ID NO: 1);

B. 5'->>>IL-12p40.CCC.TGC.AGG.GTC.CGA.TCC-GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.TCC.ATG.GGT.CAA.>>>IL-12p35-3' (SEQ ID NO: 31);

C. 5'->>>IL-12p35.5'-TAT.CTG.AGC.TCC.GCC-GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.ATG.TGG.GAG.CTG.GAG.AAA.>>>IL-12p40-3' (SEQ ID NO: 32);

D. 5'->>>IL-12p40.TGT.GTT.CCC.TGC.AGG.GTC.CGA.TCC-GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.AGG.GTC.ATT.CCA.GTC.TCT.GGA.CCT.GCC.>>>IL-12p35-3' (SEQ ID NO: 33); and

E. 5'->>>hIL-12p40.TGC.AGT.GGT.GGC.GGT.GGC.GGC.GGA.TCT.AGA.AAC.>>>hIL-12p35-3' (SEQ ID NO: 35).

No IL-12 bioactivity was detectable in media conditioned by mock-transfected CMS-5 cells, and CMS-5 cells infected with the SFG retrovirus alone, or by a related retrovirus (MFG) carrying the *lac-z* gene. However, media conditioned by these cells contained significant inhibitory activity at 1:2 and 1:10 dilutions, inhibiting as much as 95% of the bioactivity of 500 pg/ml of rmIL-12 (Table 1, and other data not shown). Despite this background of inhibitory activity in the conditioned media, bioactive IL-12 production proved to be still demonstrable.

Constructs for the expression of single subunits of the IL-12 protein (pSFG.Il-12.p35 and pSFG.IL-12.p40) resulted in no detectable bioactivity on their own. However, cotransfection of BOSC23 cells with these constructs together resulted in bioactive IL-12 secretion

by infected CMS-5 cells. Similarly, CMS-5 cells infected with the SFG.IL-12.p35 retrovirus and 24 hours later with the SFG.IL-12.p40 retrovirus also produced bioactive IL-12 (Table 1).

The dicistronic constructs designed to express both IL-12 subunits using the IRES sequence resulted in similar levels of bioactive IL-12 production (despite an undetectable level of viral infection as determined by Southern hybridization analysis (see below)) (Table 1). The ability of IRES-containing retroviruses to result in bioactive IL-12 production has been confirmed by generating stable clonal retrovirus producing cell lines using both these constructs.

All IL-12 fusion protein constructs resulted in significant bioactive IL-12 production by infected CMS-5 cells. Of particular note was the SFG.IL-12.p40 linker.Δp35 construct, for which IL-12 bioactivity was demonstrable in undiluted conditioned medium (despite the background of substantial inhibitory activity) and for which a 1:1000 dilution of conditioned medium contained bioactivity equivalent to 301 pg/ml of rmIL-12 (Table 1). All constructs resulted in titratable IL-12 bioactivity despite significant non-specific inhibitory activity in the conditioned media as well. Bioactivity of hIL-12 was confirmed at Hoffman-LaRoche (Nutley, NJ, laboratory of Dr. M. Gately), and the specific activity of the hIL-12 fusion protein was determined to be approximately equivalent to the specific activity of recombinant native human IL-12.

TABLE 1

Construct	Agonist assay (IL-12 bioactivity, pg/ml)			Antagonist assay (% inhibition of 500 pg/ml IL-12 in assay)		
	Dilution of CM in assay			Dilution of CM in assay		
	1:1	1:100	1:1000	1:2	1:10	1:1000
No DNA	<50	<50	<50	56	62	8.6
SFG-empty	<50	<50	<50	12	47	-91
MFG-lac-z	<50	<50	<50	66	56	64
SFG.IL-12p35	<50	<50	<50	65	76	45
SFG.IL-12p40	<50	<50	<50	94	84	38
2X infection[a]	199.7	234.2	137.2	1	1	-84
2X transfection[b]	244.7	118.8	<50	12	-3	-2
A	86.5	<50	<50	44	60	46
B	253.8	<50	<50	41	12	-14
C	189.2	57.0	<50	43	42	47
D	297.8	600.1	301.2	-48	-143	-93

These data are from one of three separate assays.

a. Target cells infected sequentially with pSFG.IL-12.p35 and then pSFG.IL-12.p40 viruses (each containing only the respective cDNA between the Nco1 and BamH1 sites)

b. BOSC23 cells were transfected with a mixture of pSFG.IL-12p35 and pSFG.IL-12.p40 constructs

These data indicate IL-12 agonist activity was present in media conditioned by cells infected with the fusion protein retroviral constructs. It is presumed that this results from bioactivity of secreted respective fusion proteins.

The fusion proteins were demonstrated to be present using Western blotting. Serum-free CM from wild-type CMS-5 cells or CMS-5 cells expressing native IL-12 or the IL-12 fusion protein (SFG.IL-12.p35.IRES.p40 or SFG.IL-12.p40.linker.Δp35) were collected, filtered (0.2 μm) and stored at -70°C. The CMs were concentrated 20-30-fold, and 20 μg total protein sample was run on 10% polyacrylamide gels with or without 10% β-mercaptoethanol. The primary antibody was a polyclonal goat anti-rmIL-12 antibody (gift of D. Presky, Hoffman-La Roche, NJ). The "Renaissance" detection system (NEN Dupont) was used. A preliminary analysis indicated a 4-fold greater signal resulted from CM containing the single chain IL-12 (SFG.IL-12.p40.linker.Δp35) fusion protein, and hence a 5 μg total protein sample from this CM was loaded. The control lanes were CM from wild-type cells without or spiked with 50 ng rmIL-12.

EXAMPLE 5 Southern Hybridization Analysis of Genomic DNA from Infected CMS-5 Cells

Southern hybridization analysis of genomic DNA from the populations of infected CMS-5 cells was performed to demonstrate the presence of a hybridizing band consistent with infection of these cells by retroviruses of the expected structure, and to determine the efficiency of viral infection (by determination of retroviral copy number by genome).

From these Nhe1 digests of genomic DNA, a hybridizing retrovirus-derived band of 985 bp plus the size of the insert cloned into the Nco1-BamH1 sites of SFG was

predicted (See Figure 1). The size of the various cloned fragments were: IL-12.p35 cDNA, 0.6 kb; IL-12.p40 cDNA, 1.0 kb; IRES, 0.7 kb; linker, 0.05 kb; the putative leader sequence deleted in two constructs was 0.066 bp.

The BOSC23 cell supernatants resulted in viral copy numbers of between 0.1 and 1.4 copies/genome (mostly 0.1-0.3 copies/genome) for all constructs except for the IRES-containing constructs, where no hybridizing band of the expected size (3.2 kb) was seen (Table 2).

Of particular note are the comparative results for the IL-12 fusion proteins retrovirus constructs in these populations of infected cells. Although the pSFG.IL-12.p35.linker.p40 retrovirus was present at 1.4 copies/genome, this corresponded with a relatively low level of bioactive IL-12 production (Table 2). However, the SFG.IL-12.p40 linker.Δp35 retrovirus resulted in a relatively high level of IL-12 bioactivity, although it was present at 0.2 copies/genome.

TABLE 2: Retrovirus Copy Number in CMS-5 Cells Infected by SFG.IL-12 Retroviruses

SFG.IL-12 construct containing:	Retrovirus copy number[a]
Nil	0
IL-12.p35	0.1
IL-12.p40	0.3
Sequential infection (p35/p40)	0.3/0.3
Co-transfection (p35/p40)	0.1/0.1
IL-12.p35-IRES-p40	<<0.1[b]
IL-12.p40.IRES.p35	<<0.1[b]
IL-12.p35.linker.p40	1.4[b]
IL-12.p40.linker.p35	0.1[b]
IL-12.p35.linker.Δp40	0.4[b]
IL-12.p40.linker.Δp35	0.2[b]
1 copy control	1.0[b]
0.1 copy control	0.1[b]

[a] Relative to a plasmid copy number control of 13.5 pg of pSFG.IL-12.p35 linker.p40, calculated to be equimolar to 1 copy/genome for 10 μg genomic DNA.

[b] Mean of results from one Southern blot probed first with a p35 and then with a p40 radiolabelled probe. Relative intensity of signals was quantitated using a Fuji BAS-II phosphoimager.

EXAMPLE 6 Comparison of Immunotherapy of Established Immunogenic CMS-5 Tumors with GM-CSF-secreting and IL-12-secreting Tumor Cells

Cytokine-Secreting Tumor Cells

SFG retroviruses generated by CRE or CRIP packaging cell lines were used for the transduction of tumor cells. The amount of cytokine secreted *in vitro* by the tumor cells used in these studies were (in ng/ml/48h/10^6 irradiated cells [all collected in 10 ml]): cells infected with CRIP-packaged SFG.GM-CSF, B16>250, CMS-5>250; cells infected with CRE-packaged SFG.p35.IRES.p40.IL-12, B16 1-3, CMS-5 60-400; cells infected with CRE-packaged SFG.IL-12.p40.linker.Δp35, CMS-5 490-950; cells infected with CRIP-packaged SFG.IL-12.p35.IRES.p40, B16 90; and cells infected with CRIP-packaged SFG.IL-12.p40.linker.Δp35, B16 170 and RENCA 45. The tumor cells were irradiated to prevent the formation of additional tumors therefrom after injection into mice, and cytokine secretion was characterized for the same irradiated cells. GM-CSF concentrations of conditioned media (CM) were determined by ELISA (Endogen, Cambridge) and IL-12 levels by a bioassay based on the proliferation of concanavalin-A and interleukin-2 primed splenocytes (Schoenhaut *et al*. *J. Immunol* 148:3433 (1992)).

In an initial procedure, fibrosarcoma tumors were initiated with 2 x 10^5 CMS-5 tumor cells injected subcutaneously on the back of syngeneic BALB/C mice, and immunotherapy (irradiated wild-type, GM-CSF-secreting or IL-12-secreting tumor cells) commenced either 7 or 14 days later. In this experiment, mice were stratified into multiple groups of 5 to 10 mice that received either 1, 2 or 3 weekly doses of immunotherapy at either 1 x 10^6 or 5 x 10^6 cells/dose. However, the primary analysis was

stratified only by the type of cells used as immunotherapy and the day on which treatment began, regardless of other scheduling variables.

Mice treated with irradiated IL-12-secreting tumor cells showed significantly better long-term tumor-free survival compared to untreated mice or mice treated with wild-type or GM-CSF-secreting tumor cells, for therapy schedules starting either 7 or 14 days after tumor challenge (Figures 7A and 7B, $p<0.05$ for all comparisons with IL-12-secreting tumor cell immunotherapy). When immunotherapy was commenced 7 days after tumor transplantation, much of the survival advantage of mice treated with IL-12-secreting tumor cells was due to prevention of the development of late tumors (Figure 7C, $p<0.05$ for all comparisons by the log rank test). When immunotherapy commenced 14 days after tumor transplantation, much of the survival advantage was due to the regression of established tumors (Figures 7C, 7D and 8, $p<0.005$ compared to mice receiving wild-type or GM-CSF-secreting tumor cell immunotherapy). The palpable tumors that regressed in mice receiving IL-12-secreting tumor cell immunotherapy range from 1 to 8.5 mm in average diameter, and became impalpable 20-43 days (median 30 days) after tumor transplantation. The number of animals per group were respectively: 7A and 7B, 137, 40, 40, 38; and 7C and 7D, 18, 39, 40, 40. P-values given are for the least significant difference between treatment with IL-12-secreting cells and other groups. To determine the overall effect, data were pooled from groups that received like-immunotherapy cells by several schedules.

Subgroup analyses of therapies commencing on day 14 suggested that superior survival from immunotherapy with IL-12-secreting tumor cells resulted from schedules with more than one weekly dose of immunotherapy ($p=0.1$), doses of 5×10^6 rather than 1×10^6 IL-12-secreting cells

($p<0.02$). Hence, in all subsequent experiments utilizing transduced tumor cells, immunotherapy regimens comprised the higher cell dose administered weekly for 4 weeks.

Statistical Analyses

All analyses were conducted on the basis of the intention to treat at the time of the random allocation of mice to groups. Descriptive statistics were calculated for major endpoints. Except where otherwise stated, differences in the survival endpoint were evaluated by the Wilcoxon rank-sum test. For survival analyses, occasional deaths immediately after anaesthesia and treatment were treated as censored events. The chi-square test was used to measure the association of categorical variables. Where p-values summarize the comparisons between multiple groups, only the largest p-value is given. Analyses were conducted using JMP software on a Power Macintosh 6100/60 computer.

EXAMPLE 7 Study of Mechanisms of IL-12-induced Tumor Regression and Improved Survival

In order to determine whether immunotherapy with IL-12-secreting tumor cells was effective against larger tumor burdens, tumors were established with 4×10^5 tumor cells, which, compared with 2×10^5 cells, resulted in higher tumor incidence (day 14 palpability rates of 98/100 vs. 83/100, respectively), larger mean tumor size (6.7 ± 3.0 vs. 3.7 ± 2.4 mm diameter at day 14, respectively) and shorter median survival without treatment (31 vs. 37 days). Following establishment of these larger tumors, 70% (7/10) of tumor-bearing mice treated with IL-12-secreting tumor cell immunotherapy from day 14 survived with complete tumor regression, compared with 0/10 mice treated with wild-type tumor cells ($p \leq 0.001$).

The administration of IL-12 systemically (intraperitoneally) to mice bearing tumors initiated by $2 \times$

10^5 CMS-5 cells also resulted in the regression of established tumors (Figure 9) and improved survival (4/5 at 90 days for mice treated with 0.1 μg/d, compared to 0/5 for placebo-treated mice). For mice with tumors established from 2 x 10^5 CMS-5 cells, an IL-12 dose of 0.1 μg/d (4/5 survival) was superior to 1 (3/5 survival), 0.01 (1/5 survival) and 0.001 μg/d, for regimens starting either 7 or 14 days after tumor transplantation.

It was therefore possible that the regression of tumors in mice receiving immunotherapy with IL-12-secreting cells was not dependent on the local release of IL-12 at the site of the irradiated tumor cells administered as immunotherapy, but rather on a systemic effect of IL-12. This was evaluated by comparing different schedules starting on day 14 which combined either wild-type or IL-12-secreting tumor cell immunotherapy and systemic therapy with either IL-12, placebo or nothing. In mice with tumors initiated by 2 x 10^5 CMS-5 cells, there was a tendency for median and overall survival to be better for immunotherapy with IL-12-secreting tumor cells than for systemic IL-12 therapy (Figure 10A). This tendency was statistically significant for mice with tumors initiated by 4 x 10^5 cells (Figure 10B, p=0.006 comparing groups receiving systemic IL-12 (either alone or in combination with wild-type cells) vs. mice receiving IL-12-secreting tumor cell immunotherapy (either alone or with systemic therapy with diluent)). Comparisons of smaller, uniformly-treated groups (n=10/group, tumors initiated by 4 x 10^5 cells) indicated that combining administration of wild type cells with systemic IL-12 was not different to systemic IL-12 therapy alone (p=0.85) and appeared inferior to vaccination with IL-12-secreting tumor cells alone (p=0.04) or given with placebo systemic therapy (p=0.19).

EXAMPLE 8 Antitumor Effect of IL-12 Fusion Protein

In the pre-existing CMS-5 tumor model, immunotherapy with CMS-5 tumor cells expressing the IL-12 fusion protein SFG.IL-12.p40.linker.Δp35 was as effective as therapy with tumor cells making native IL-12 (Figures 11A and 11B). For mice with tumors initiated by either 2 x 10^5 or 4 x 10^5 CMS-5, survival was greater than 90% for groups of mice treated with CMS-5 cells secreting either form of IL-12, compared to less than 40% for mice that received no treatment, or treatment with wild-type or GM-CSF-secreting cells ($p \leq 0.02$).

EXAMPLE 9 Comparison of Immunotherapy of Established Non-immunogenic B16 Tumors with GM-CSF-secreting and IL-12-secreting Tumor Cells

In order to assess the efficacy of immunotherapy with IL-12-secreting tumor cells in another tumor model, the non-immunogenic B16 melanoma was studied. B16 tumor cells were transduced to make native IL-12 at 90 ng/ml/48hr/10^6 irradiated cells or the single chain IL-12 (SFG.IL-12.p40.linker.Δp35) at 170 ng/ml/48hr/10^6 irradiated cells. B16 tumors were initiated with 4 x 10^5 cells and immunotherapy of established tumors commenced either on day 7 (25% tumor palpability) or day 14 (93% tumor palpability, mean tumor diameter 5.74±3.23, n=56)). This procedure was analyzed after 31 days of follow up, when only 1/60 (2%) of mice that were treated with wild-type cells, CM-CSF-secreting cells or nothing as immunotherapy survived. Although mice treated with IL-12-secreting cells had comparably poor overall survival, their median survival was significantly prolonged compared to that of control mice treated with wild-type cells when treatment commenced on day 7 (Figure 12A, 24 vs. 18 days $p=0.01$) and day 14 (Figure 12B, 28 vs. 18 days, $p=0.0005$). Similarly, median survival was prolonged with therapy with IL-12 fusion

protein-secreting tumor cells when treatment commenced on day 7 (21 vs. 18 days, p=0.08) and day 14 (24 vs. 18 days, p=0.006). In 3/4 scenarios, IL-12-secreting tumor cells were superior to GM-CSF-secreting cells (respective p-values 0.01, 0.14, 0.003, 0.02).

Given the potent effect of GM-CSF-secreting B16 cells to induce antitumor immunity when used as a vaccine prior to tumor challenge, but their lack of effect on tumor growth when administered after tumor establishment, the effects of IL-12-secreting and GM-CSF-secreting B16 cells were compared as vaccines in a B16 tumor challenge model. The IL-12-secreting B16 cells used in initial studies secreted native IL-12 at 1-3 ng/ml/48hr/10^6 irradiated cells. GM-CSF-secreting B16 cells induced antitumor immunity when used as vaccines before tumor transplantation (Figure 12C, 80% 100-day survival).

EXAMPLE 10 Immunotherapeutic Effect of IL-12 Delivery by Tumor Cells of Different Origin from Tumor to be Treated

The effect of the delivery of IL-12 by tumor cells of different origin from the tumor to be treated on survival was assessed in renal cell carcinoma (RENCA) tumors. RENCA tumors were initiated with 4 x 10^5 cells, and immunotherapy of established tumors commenced on day 14. In one procedure, groups of mice were treated with either irradiated wild type CMS-5 cells or CMS-5 tumor cells transduced to secrete either native IL-12 or the fusion protein SFG.IL-12.p40.linker.Δp35 (Figure 13A). In another procedure, additional groups of mice were treated with a combination of irradiated CMS-5 and RENCA wild type cells, a combination of irradiated wild type CMS-5 and IL-12-secreting RENCA tumor cells or a combination of irradiated wild type RENCA tumor cells and IL-12-secreting CMS-5 tumor cells (Figure 13B).

In the first procedure, immunotherapy with both CMS-5 tumor cells secreting native IL-12 and the IL-12 fusion protein prolonged the median survival (p-values of p=0.02 and p=0.06, respectively). In the second procedure, mice treated with a combination of irradiated RENCA tumor cells and IL-12-secreting CMS-5 tumor cells exhibited a trend toward increased survival.

Additionally, CMS-5 tumors were initiated with 4×10^5 cells, and immunotherapy of established tumors commenced on day 14. In one procedure, the established CMS-5 tumors were treated with either irradiated wild type RENCA tumor cells or RENCA tumor cells transduced to secrete either native IL-12 or the IL-12 fusion protein SFG.IL-12.p40.linker.Δp35 (Figure 14A). In another procedure, additional groups of mice were treated with a combination of irradiated CMS-5 and RENCA wild type cells, a combination of irradiated wild type CMS-5 and IL-12-secreting RENCA tumor cells or a combination of irradiated wild type RENCA tumor cells and IL-12-secreting CMS-5 tumor cells (Figure 14B).

In the first procedure, immunotherapy with RENCA cells secreting the IL-12 fusion protein moderately prolonged the survival of the mice compared to mice treated with wild type RENCA cells; the effect was consistent with the lower dose of IL-12 delivered by the transduced RENCA cells. In the second procedure, both mice treated with a combination of IL-12-secreting CMS-5 cells and wild type RENCA cells and mice treated with a combination of IL-12-secreting CMS-5 cells and wild type RENCA cells showed significantly prolonged survival (p-values of p=0.004 and p=0.04) compared with mice treated with wild type RENCA and wild type CMS-5 cells.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Whitehead Institute for Biomedical Research
(B) STREET: 9 Cambridge Center
(C) CITY: Cambridge
(D) STATE/PROVINCE: MA
(E) COUNTRY: US
(F) POSTAL CODE/ZIP: 02142
(G) TELEPHONE: 617-258-5104
(I) TELEFAX: 617-258-6294

(i) APPLICANT/INVENTOR:
(A) NAME: Graham J. Lieschke
(B) STREET: 5 Rollins Court
(C) CITY: Cambridge
(D) STATE/PROVINCE: MA
(E) COUNTRY: US
(F) POSTAL CODE/ZIP: 02139

(i) APPLICANT/INVENTOR:
(A) NAME: Richard C. Mulligan
(B) STREET: 2 Sandy Pond Road
(C) CITY: Lincoln
(D) STATE/PROVINCE: MA
(E) COUNTRY: US
(F) POSTAL CODE/ZIP: 01773

(ii) TITLE OF INVENTION: Bioactive Fusion Proteins and Pre-existing Tumor Therapy

(iii) NUMBER OF SEQUENCES: 36

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
(B) STREET: Two Militia Drive

(C) CITY: Lexington
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02173

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/385,335
(B) FILING DATE: 08-FEB-1995

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Granahan, Patricia
(B) REGISTRATION NUMBER: 32,227
(C) REFERENCE/DOCKET NUMBER: WHI95-01A.PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 617-861-6240
(B) TELEFAX: 617-861-9540

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGCTCCGCCG GTGGTGGTGG GTCGGGTGGC GGCGGATCTT CCATGGGTCC TCAG 54

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTCCGATCCG GTGGCGGTGG CTCGGGCGGT GGTGGGTCGG GTGGCGGCGG ATCTTCCATG 60

GGTCAA 66

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCTCCGCCG GTGGCGGTGG CTCGGGCGGT GGTGGGTCGG GTGGCGGCGG ATCTATGTGG 60

GAGCTG 66

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTCCGATCCG GTGGCGGTGG CTCGGGCGGT GGTGGGTCGG GTGGCGGCGG ATCTAGGGTC 60

ATTCCA 66

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ser
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6350 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AAGCTTTGCT CTTAGGAGTT TCCTAATACA TCCCAAACTC AAATATATAA AGCATTTGAC 60

TTGTTCTATG CCCTAGGGGG CGGGGGGAAG CTAAGCCAGC TTTTTTTAAC ATTTAAAATG 120

TTAATTCCAT TTTAAATGCA CAGATGTTTT TATTTCATAA GGGTTTCAAT GTGCATGAAT 180

GCTGCAATAT TCCTGTTACC AAAGCTAGTA TAAATAAAAA TAGATAAACG TGGAAATTAC 240

TTAGAGTTTC TGTCATTAAC GTTTCCTTCC TCAGTTGACA ACATAAATGC GCTGCTGAGC 300

AAGCCAGTTT GCATCTGTCA GGATCAATTT CCCATTATGC CAGTCATATT AATTACTAGT 360

CAATTAGTTG ATTTTTATTT TTGACATATA CATGTGAATG AAAGACCCCA CCTGTAGGTT 420

TGGCAAGCTA GCTTAAGTAA CGCCATTTTG CAAGGCATGG AAAAATACAT AACTGAGAAT 480

AGAAAAGTTC AGATCAAGGT CAGGAACAGA TGGAACAGCT GAATATGGGC CAAACAGGAT 540

ATCTGTGGTA AGCAGTTCCT GCCCCGGCTC AGGGCCAAGA ACAGATGGAA CAGCTGAATA 600

TGGGCCAAAC AGGATATCTG TGGTAAGCAG TTCCTGCCCC GGCTCAGGGC CAAGAACAGA 660

TGGTCCCCAG ATGCGGTCCA GCCCTCAGCA GTTTCTAGAG AACCATCAGA TGTTTCCAGG 720

GTGCCCCAAG GACCTGAAAT GACCCTGTGC CTTATTTGAA CTAACCAATC AGTTCGCTTC 780

TCGCTTCTGT TCGCGCGCTT ATGCTCCCCG AGCTCAATAA AAGAGCCCAC AACCCCTCAC 840

TCGGGGCGCC AGTCCTCCGA TTGACTGAGT CGCCCGGGTA CCCGTGTATC CAATAAACCC 900

TCTTGCAGTT GCATCCGACT TGTGGTCTCG CTGTTCCTTG GGAGGGTCTC CTCTGAGTGA 960

TTGACTACCC GTCAGCGGGG GTCTTTCATT TGGGGGCTCG TCCGGGATCG GGAGACCCCT 1020

GCCCAGGGAC CACCGACCCA CCACCGGGAG GTAAGCTGGC CAGCAACTTA TCTGTGTCTG 1080

TCCGATTGTC TAGTGTCTAT GACTGATTTT ATGCGCCTGC GTCGGTACTA GTTAGCTAAC 1140

TAGCTCTGTA TCTGGCGGAC CCGTGGTGGA ACTGACGAGT TCGGAACACC CGGCCGCAAC 1200

CCTGGGAGAC GTCCCAGGGA CTTCGGGGGC CGTTTTTGTG GCCCGACCTG AGTCCTAAAA 1260

TCCCGATCGT TTAGGACTCT TTGGTGCACC CCCCTTAGAG GAGGGATATG TGGTTCTGGT 1320

AGGAGACGAG AACCTAAAAC AGTTCCCGCC TCCGTCTGAA TTTTTGCTTT CGGTTTGGGA 1380

CCGAAGCCGC GCCGCGCGTC TTGTCTGCTG CAGCATCGTT CTGTGTTGTC TCTGTCTGAC 1440

TGTGTTTCTG TATTTGTCTG AAAATATGGG CCCGGGCTAG CCTGTTACCA CTCCCTTAAG 1500

TTTGACCTTA GGTCACTGGA AAGATGTCGA GCGGATCGCT CACAACCAGT CGGTAGATGT 1560

CAAGAAGAGA CGTTGGGTTA CCTTCTGCTC TGCAGAATGG CCAACCTTTA ACGTCGGATG 1620

GCCGCGAGAC GGCACCTTTA ACCGAGACCT CATCACCCAG GTTAAGATCA AGGTCTTTTC 1680

ACCTGGCCCG CATGGACACC CAGACCAGGT GGGGTACATC GTGACCTGGG AAGCCTTGGC 1740

TTTTGACCCC CCTCCCTGGG TCAAGCCCTT TGTACACCCT AAGCCTCCGC CTCCTCTTCC 1800

TCCATCCGCC CCGTCTCTCC CCCTTGAACC TCCTCGTTCG ACCCCGCCTC GATCCTCCCT 1860

TTATCCAGCC CTCACTCCTT CTCTAGGCGC CCCCATATGG CCATATGAGA TCTTATATGG 1920

GGCACCCCCG CCCCTTGTAA ACTTCCCTGA CCCTGACATG ACAAGAGTTA CTAACAGCCC 1980

CTCTCTCCAA GCTCACTTAC AGGCTCTCTA CTTAGTCCAG CACGAAGTCT GGAGACCTCT 2040

GGCGGCAGCC TACCAAGAAC AACTGGACCG ACCGGTGGTA CCTCACCCTT ACCGAGTCGG 2100

CGACACAGTG TGGGTCCGCC GACACCAGAC TAAGAACCTA GAACCTCGCT GGAAAGGACC 2160

TTACACAGTC CTGCTGACCA CCCCCACCGC CCTCAAAGTA GACGGCATCG CAGCTTGGAT 2220

ACACGCCGCC CACGTGAAGG CTGCCGACCC CGGGGGTGGA CCATCCTCTA GACTGCCATG 2280

GCGCGGATCC GGATTAGTCC AATTTGTTAA AGACAGGATA TCAGTGGTCC AGGCTCTAGT 2340

TTTGACTCAA CAATATCACC AGCTGAAGCC TATAGAGTAC GAGCCATAGA TAAAATAAAA 2400

GATTTTATTT AGTCTCCAGA AAAAGGGGGG AATGAAAGAC CCCACCTGTA GGTTTGGCAA 2460

GCTAGCTTAA GTAACGCCAT TTTGCAAGGC ATGGAAAAAT ACATAACTGA GAATAGAGAA 2520

GTTCAGATCA AGGTCAGGAA CAGATGGAAC AGCTGAATAT GGGCCAAACA GGATATCTGT 2580

GGTAAGCAGT TCCTGCCCCG GCTCAGGGCC AAGAACAGAT GGAACAGCTG AATATGGGCC 2640

AAACAGGATA TCTGTGGTAA GCAGTTCCTG CCCCGGCTCA GGGCCAAGAA CAGATGGTCC 2700

CCAGATGCGG TCCAGCCCTC AGCAGTTTCT AGAGAACCAT CAGATGTTTC CAGGGTGCCC 2760

CAAGGACCTG AAATGACCCT GTGCCTTATT TGAACTAACC AATCAGTTCG CTTCTCGCTT 2820

CTGTTCGCGC GCTTCTGCTC CCCGAGCTCA ATAAAAGAGC CCACAACCCC TCACTCGGGG 2880

CGCCAGTCCT CCGATTGACT GAGTCGCCCG GGTACCCGTG TATCCAATAA ACCCTCTTGC 2940

AGTTGCATCC GACTTGTGGT CTCGCTGTTC CTTGGGAGGG TCTCCTCTGA GTGATTGACT 3000

ACCCGTCAGC GGGGGTCTTT CACACATGCA GCATGTATCA AAATTAATTT GGTTTTTTTT 3060

CTTAAGTATT TACATTAAAT GGCCATAGTA CTTAAAGTTA CATTGGCTTC CTTGAAATAA 3120

ACATGGAGTA TTCAGAATGT GTCATAAATA TTTCTAATTT TAAGATAGTA TCTCCATTGG 3180

CTTTCTACTT TTTCTTTTAT TTTTTTTTGT CCTCTGTCTT CCATTTGTTG TTGTTGTTGT 3240

TTGTTTGTTT GTTTGTTGGT TGGTTGGTTA ATTTTTTTTT AAAGATCCTA CACTATAGTT 3300

CAAGCTAGAC TATTAGCTAC TCTGTAACCC AGGGTGACCT TGAAGTCATG GGTAGCCTGC 3360

TGTTTTAGCC TTCCCACATC TAAGATTACA GGTATGAGCT ATCATTTTTG GTATATTGAT 3420

TGATTGATTG ATTGATGTGT GTGTGTGTGA TTGTGTTTGT GTGTGTGACT GTGAAAATGT 3480

GTGTATGGGT GTGTGTGAAT GTGTGTATGT ATGTGTGTGT GTGAGTGTGT GTGTGTGTGT 3540

GTGCATGTGT GTGTGTGTGA CTGTGTCTAT GTGTATGACT GTGTGTGTGT GTGTGTGTGT 3600

GTGTGTGTGT GTGTGTGTGT GTGTGTTGTG AAAAAATATT CTATGGTAGT GAGAGCCAAC 3660

GCTCCGGCTC AGGTGTCAGG TTGGTTTTTG AGACAGAGTC TTTCACTTAG CTTGGAATTC 3720

ACTGGCCGTC GTTTTACAAC GTCGTGACTG GGAAAACCCT GGCGTTACCC AACTTAATCG 3780

CCTTGCAGCA CATCCCCCTT TCGCCAGCTG GCGTAATAGC GAAGAGGCCC GCACCGATCG 3840

CCCTTCCCAA CAGTTGCGCA GCCTGAATGG CGAATGGCGC CTGATGCGGT ATTTTCTCCT 3900

TACGCATCTG TGCGGTATTT CACACCGCAT ATGGTGCACT CTCAGTACAA TCTGCTCTGA 3960

TGCCGCATAG TTAAGCCAGC CCCGACACCC GCCAACACCC GCTGACGCGC CCTGACGGGC 4020

TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA GCTGCATGTG 4080

TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGATGACG AAAGGGCCTC GTGATACGCC 4140

TATTTTTATA GGTTAATGTC ATGATAATAA TGGTTTCTTA GACGTCAGGT GGCACTTTTC 4200

GGGGAAATGT GCGCGGAACC CCTATTTGTT TATTTTTCTA AATACATTCA AATATGTATC 4260

CGCTCATGAG ACAATAACCC TGATAAATGC TTCAATAATA TTGAAAAAGG AAGAGTATGA 4320

GTATTCAACA TTTCCGTGTC GCCCTTATTC CCTTTTTTGC GGCATTTTGC CTTCCTGTTT 4380

TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA AGATCAGTTG GGTGCACGAG 4440

TGGGTTACAT CGAACTGGAT CTCAACAGCG GTAAGATCCT TGAGAGTTTT CGCCCCGAAG 4500

AACGTTTTCC AATGATGAGC ACTTTTAAAG TTCTGCTATG TGGCGCGGTA TTATCCCGTA 4560

TTGACGCCGG GCAAGAGCAA CTCGGTCGCC GCATACACTA TTCTCAGAAT GACTTGGTTG 4620

AGTACTCACC AGTCACAGAA AAGCATCTTA CGGATGGCAT GACAGTAAGA GAATTATGCA 4680

GTGCTGCCAT AACCATGAGT GATAACACTG CGGCCAACTT ACTTCTGACA ACGATCGGAG 4740

GACCGAAGGA GCTAACCGCT TTTTTGCACA ACATGGGGGA TCATGTAACT CGCCTTGATC 4800

GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA GCGTGACACC ACGATGCCTG 4860

TAGCAATGGC AACAACGTTG CGCAAACTAT TAACTGGCGA ACTACTTACT CTAGCTTCCC 4920

GGCAACAATT AATAGACTGG ATGGAGGCGG ATAAAGTTGC AGGACCACTT CTGCGCTCGG 4980

CCCTTCCGGC TGGCTGGTTT ATTGCTGATA AATCTGGAGC CGGTGAGCGT GGGTCTCGCG 5040

GTATCATTGC AGCACTGGGG CCAGATGGTA AGCCCTCCCG TATCGTAGTT ATCTACACGA 5100

CGGGGAGTCA GGCAACTATG GATGAACGAA ATAGACAGAT CGCTGAGATA GGTGCCTCAC 5160

TGATTAAGCA TTGGTAACTG TCAGACCAAG TTTACTCATA TATACTTTAG ATTGATTTAA 5220

AACTTCATTT TTAATTTAAA AGGATCTAGG TGAAGATCCT TTTTGATAAT CTCATGACCA 5280

AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGCGTCAGA CCCCGTAGAA AAGATCAAAG 5340

GATCTTCTTG AGATCCTTTT TTTCTGCGCG TAATCTGCTG CTTGCAAACA AAAAAACCAC 5400

CGCTACCAGC GGTGGTTTGT TTGCCGGATC AAGAGCTACC AACTCTTTTT CCGAAGGTAA 5460

CTGGCTTCAG CAGAGCGCAG ATACCAAATA CTGTCCTTCT AGTGTAGCCG TAGTTAGGCC 5520

ACCACTTCAA GAACTCTGTA GCACCGCCTA CATACCTCGC TCTGCTAATC CTGTTACCAG 5580

TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGACTCAAGA CGATAGTTAC 5640

CGGATAAGGC GCAGCGGTCG GGCTGAACGG GGGGTTCGTG CACACAGCCC AGCTTGGAGC 5700

GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCA TTGAGAAAGC GCCACGCTTC 5760

CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTCGGAACA GGAGAGCGCA 5820

CGAGGGAGCT TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG TTTCGCCACC 5880

TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG GCGGAGCCTA TGGAAAAACG 5940

CCAGCAACGC GGCCTTTTTA CGGTTCCTGG CCTTTTGCTG GCCTTTTGCT CACATGTTCT 6000

TTCCTGCGTT ATCCCCTGAT TCTGTGGATA ACCGTATTAC CGCCTTTGAG TGAGCTGATA 6060

CCGCTCGCCG CAGCCGAACG ACCGAGCGCA GCGAGTCAGT GAGCGAGGAA GCGGAAGAGC 6120

GCCCAATACG CAAACCGCCT CTCCCCGCGC GTTGGCCGAT TCATTAATGC AGCTGGCACG 6180

ACAGGTTTCC CGACTGGAAA GCGGGCAGTG AGCGCAACGC AATTAATGTG AGTTAGCTCA 6240

CTCATTAGGC ACCCCAGGCT TTACACTTTA TGCTTCCGGC TCGTATGTTG TGTGGAATTG 6300

TGAGCGGATA ACAATTTCAC ACAGGAAACA GCTATGACCA TGATTACGCC 6350

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6350 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TTCGAAACGA GAATCCTCAA AGGATTATGT AGGGTTTGAG TTTATATATT TCGTAAACTG 60

AACAAGATAC GGGATCCCCC GCCCCCCTTC GATTCGGTCG AAAAAAATTG TAAATTTTAC 120

AATTAAGGTA AAATTTACGT GTCTACAAAA ATAAAGTATT CCCAAAGTTA CACGTACTTA 180

CGACGTTATA AGGACAATGG TTTCGATCAT ATTTATTTTT ATCTATTTGC ACCTTTAATG 240

AATCTCAAAG ACAGTAATTG CAAAGGAAGG AGTCAACTGT TGTATTTACG CGACGACTCG 300

TTCGGTCAAA CGTAGACAGT CCTAGTTAAA GGGTAATACG GTCAGTATAA TTAATGATCA 360

GTTAATCAAC TAAAAATAAA AACTGTATAT GTACACTTAC TTTCTGGGGT GGACATCCAA 420

ACCGTTCGAT CGAATTCATT GCGGTAAAAC GTTCCGTACC TTTTTATGTA TTGACTCTTA 480

TCTTTTCAAG TCTAGTTCCA GTCCTTGTCT ACCTTGTCGA CTTATACCCG GTTTGTCCTA 540

TAGACACCAT TCGTCAAGGA CGGGGCCGAG TCCCGGTTCT TGTCTACCTT GTCGACTTAT 600

ACCCGGTTTG TCCTATAGAC ACCATTCGTC AAGGACGGGG CCGAGTCCCG GTTCTTGTCT 660

ACCAGGGGTC TACGCCAGGT CGGGAGTCGT CAAAGATCTC TTGGTAGTCT ACAAAGGTCC 720

CACGGGGTTC CTGGACTTTA CTGGGACACG GAATAAACTT GATTGGTTAG TCAAGCGAAG 780

AGCGAAGACA AGCGCGCGAA TACGAGGGGC TCGAGTTATT TTCTCGGGTG TTGGGGAGTG 840

AGCCCCGCGG TCAGGAGGCT AACTGACTCA GCGGGCCCAT GGGCACATAG GTTATTTGGG 900

AGAACGTCAA CGTAGGCTGA ACACCAGAGC GACAAGGAAC CCTCCCAGAG GAGACTCACT 960

AACTGATGGG CAGTCGCCCC CAGAAAGTAA ACCCCCGAGC AGGCCCTAGC CCTCTGGGGA 1020

CGGGTCCCTG GTGGCTGGGT GGTGGCCCTC CATTCGACCG GTCGTTGAAT AGACACAGAC 1080

AGGCTAACAG ATCACAGATA CTGACTAAAA TACGCGGACG CAGCCATGAT CAATCGATTG 1140

ATCGAGACAT AGACCGCCTG GGCACCACCT TGACTGCTCA AGCCTTGTGG GCCGGCGTTG 1200

GGACCCTCTG CAGGGTCCCT GAAGCCCCCG GCAAAAACAC CGGGCTGGAC TCAGGATTTT 1260

AGGGCTAGCA AATCCTGAGA AACCACGTGG GGGGAATCTC CTCCCTATAC ACCAAGACCA 1320

TCCTCTGCTC TTGGATTTTG TCAAGGGCGG AGGCAGACTT AAAAACGAAA GCCAAACCCT 1380

GGCTTCGGCG CGGCGCGCAG AACAGACGAC GTCGTAGCAA GACACAACAG AGACAGACTG 1440

ACACAAAGAC ATAAACAGAC TTTTATACCC GGGCCCGATC GGACAATGGT GAGGGAATTC 1500

AAACTGGAAT CCAGTGACCT TTCTACAGCT CGCCTAGCGA GTGTTGGTCA GCCATCTACA 1560

GTTCTTCTCT GCAACCCAAT GGAAGACGAG ACGTCTTACC GGTTGGAAAT TGCAGCCTAC 1620

CGGCGCTCTG CCGTGGAAAT TGGCTCTGGA GTAGTGGGTC CAATTCTAGT TCCAGAAAAG 1680

TGGACCGGGC GTACCTGTGG GTCTGGTCCA CCCCATGTAG CACTGGACCC TTCGGAACCG 1740

AAAACTGGGG GGAGGGACCC AGTTCGGGAA ACATGTGGGA TTCGGAGGCG GAGGAGAAGG 1800

AGGTAGGCGG GGCAGAGAGG GGGAACTTGG AGGAGCAAGC TGGGGCGGAG CTAGGAGGGA 1860

AATAGGTCGG GAGTGAGGAA GAGATCCGCG GGGGTATACC GGTATACTCT AGAATATACC 1920

CCGTGGGGGC GGGGAACATT TGAAGGGACT GGGACTGTAC TGTTCTCAAT GATTGTCGGG 1980

GAGAGAGGTT CGAGTGAATG TCCGAGAGAT GAATCAGGTC GTGCTTCAGA CCTCTGGAGA 2040

CCGCCGTCGG ATGGTTCTTG TTGACCTGGC TGGCCACCAT GGAGTGGGAA TGGCTCAGCC 2100

GCTGTGTCAC ACCCAGGCGG CTGTGGTCTG ATTCTTGGAT CTTGGAGCGA CCTTTCCTGG 2160

AATGTGTCAG GACGACTGGT GGGGGTGGCG GGAGTTTCAT CTGCCGTAGC GTCGAACCTA 2220

TGTGCGGCGG GTGCACTTCC GACGGCTGGG GCCCCCACCT GGTAGGAGAT CTGACGGTAC 2280

CGCGCCTAGG CCTAATCAGG TTAAACAATT TCTGTCCTAT AGTCACCAGG TCCGAGATCA 2340

AAACTGAGTT GTTATAGTGG TCGACTTCGG ATATCTCATG CTCGGTATCT ATTTTATTTT 2400

CTAAAATAAA TCAGAGGTCT TTTTCCCCCC TTACTTTCTG GGGTGGACAT CCAAACCGTT 2460

CGATCGAATT CATTGCGGTA AAACGTTCCG TACCTTTTTA TGTATTGACT CTTATCTCTT 2520

CAAGTCTAGT TCCAGTCCTT GTCTACCTTG TCGACTTATA CCCGGTTTGT CCTATAGACA 2580

CCATTCGTCA AGGACGGGGC CGAGTCCCGG TTCTTGTCTA CCTTGTCGAC TTATACCCGG 2640

TTTGTCCTAT AGACACCATT CGTCAAGGAC GGGGCCGAGT CCCGGTTCTT GTCTACCAGG 2700

GGTCTACGCC AGGTCGGGAG TCGTCAAAGA TCTCTTGGTA GTCTACAAAG GTCCCACGGG 2760

GTTCCTGGAC TTTACTGGGA CACGGAATAA ACTTGATTGG TTAGTCAAGC GAAGAGCGAA 2820

GACAAGCGCG CGAAGACGAG GGGCTCGAGT TATTTTCTCG GGTGTTGGGG AGTGAGCCCC 2880

GCGGTCAGGA GGCTAACTGA CTCAGCGGGC CCATGGGCAC ATAGGTTATT TGGGAGAACG 2940

TCAACGTAGG CTGAACACCA GAGCGACAAG GAACCCTCCC AGAGGAGACT CACTAACTGA 3000

TGGGCAGTCG CCCCCAGAAA GTGTGTACGT CGTACATAGT TTTAATTAAA CCAAAAAAAA 3060

GAATTCATAA ATGTAATTTA CCGGTATCAT GAATTTCAAT GTAACCGAAG GAACTTTATT 3120

TGTACCTCAT AAGTCTTACA CAGTATTTAT AAAGATTAAA ATTCTATCAT AGAGGTAACC 3180

GAAAGATGAA AAAGAAAATA AAAAAAAACA GGAGACAGAA GGTAAACAAC AACAACAACA 3240

AACAAACAAA CAAACAACCA ACCAACCAAT TAAAAAAAAA TTTCTAGGAT GTGATATCAA 3300

GTTCGATCTG ATAATCGATG AGACATTGGG TCCCACTGGA ACTTCAGTAC CCATCGGACG 3360

ACAAAATCGG AAGGGTGTAG ATTCTAATGT CCATACTCGA TAGTAAAAAC CATATAACTA 3420

ACTAACTAAC TAACTACACA CACACACACT AACACAAACA CACACACTGA CACTTTTACA 3480

CACATACCCA CACACACTTA CACACATACA TACACACACA CACTCACACA CACACACACA 3540

CACGTACACA CACACACACT GACACAGATA CACATACTGA CACACACACA CACACACACA 3600

CACACACACA CACACACACA CACACAACAC TTTTTTATAA GATACCATCA CTCTCGGTTG 3660

CGAGGCCGAG TCCACAGTCC AACCAAAAAC TCTGTCTCAG AAAGTGAATC GAACCTTAAG 3720

TGACCGGCAG CAAAATGTTG CAGCACTGAC CCTTTTGGGA CCGCAATGGG TTGAATTAGC 3780

GGAACGTCGT GTAGGGGGAA AGCGGTCGAC CGCATTATCG CTTCTCCGGG CGTGGCTAGC 3840

GGGAAGGGTT GTCAACGCGT CGGACTTACC GCTTACCGCG GACTACGCCA TAAAAGAGGA 3900

ATGCGTAGAC ACGCCATAAA GTGTGGCGTA TACCACGTGA GAGTCATGTT AGACGAGACT 3960

ACGGCGTATC AATTCGGTCG GGGCTGTGGG CGGTTGTGGG CGACTGCGCG GGACTGCCCG 4020

AACAGACGAG GGCCGTAGGC GAATGTCTGT TCGACACTGG CAGAGGCCCT CGACGTACAC 4080

AGTCTCCAAA AGTGGCAGTA GTGGCTTTGC GCGCTACTGC TTTCCCGGAG CACTATGCGG 4140

ATAAAAATAT CCAATTACAG TACTATTATT ACCAAAGAAT CTGCAGTCCA CCGTGAAAAG 4200

CCCCTTTACA CGCGCCTTGG GGATAAACAA ATAAAAAGAT TTATGTAAGT TTATACATAG 4260

GCGAGTACTC TGTTATTGGG ACTATTTACG AAGTTATTAT AACTTTTTCC TTCTCATACT 4320

CATAAGTTGT AAAGGCACAG CGGGAATAAG GGAAAAAACG CCGTAAAACG GAAGGACAAA 4380

AACGAGTGGG TCTTTGCGAC CACTTTCATT TTCTACGACT TCTAGTCAAC CCACGTGCTC 4440

ACCCAATGTA GCTTGACCTA GAGTTGTCGC CATTCTAGGA ACTCTCAAAA GCGGGGCTTC 4500

TTGCAAAAGG TTACTACTCG TGAAAATTTC AAGACGATAC ACCGCGCCAT AATAGGGCAT 4560

AACTGCGGCC CGTTCTCGTT GAGCCAGCGG CGTATGTGAT AAGAGTCTTA CTGAACCAAC 4620

TCATGAGTGG TCAGTGTCTT TTCGTAGAAT GCCTACCGTA CTGTCATTCT CTTAATACGT 4680

CACGACGGTA TTGGTACTCA CTATTGTGAC GCCGGTTGAA TGAAGACTGT TGCTAGCCTC 4740

CTGGCTTCCT CGATTGGCGA AAAAACGTGT TGTACCCCCT AGTACATTGA GCGGAACTAG 4800

CAACCCTTGG CCTCGACTTA CTTCGGTATG GTTTGCTGCT CGCACTGTGG TGCTACGGAC 4860

ATCGTTACCG TTGTTGCAAC GCGTTTGATA ATTGACCGCT TGATGAATGA GATCGAAGGG 4920

CCGTTGTTAA TTATCTGACC TACCTCCGCC TATTTCAACG TCCTGGTGAA GACGCGAGCC 4980

GGGAAGGCCG ACCGACCAAA TAACGACTAT TTAGACCTCG GCCACTCGCA CCCAGAGCGC 5040

CATAGTAACG TCGTGACCCC GGTCTACCAT TCGGGAGGGC ATAGCATCAA TAGATGTGCT 5100

GCCCCTCAGT CCGTTGATAC CTACTTGCTT TATCTGTCTA GCGACTCTAT CCACGGAGTG 5160

ACTAATTCGT AACCATTGAC AGTCTGGTTC AAATGAGTAT ATATGAAATC TAACTAAATT 5220

TTGAAGTAAA AATTAAATTT TCCTAGATCC ACTTCTAGGA AAAACTATTA GAGTACTGGT 5280

TTTAGGGAAT TGCACTCAAA AGCAAGGTGA CTCGCAGTCT GGGGCATCTT TTCTAGTTTC 5340

CTAGAAGAAC TCTAGGAAAA AAAGACGCGC ATTAGACGAC GAACGTTTGT TTTTTTGGTG 5400

GCGATGGTCG CCACCAAACA AACGGCCTAG TTCTCGATGG TTGAGAAAAA GGCTTCCATT 5460

GACCGAAGTC GTCTCGCGTC TATGGTTTAT GACAGGAAGA TCACATCGGC ATCAATCCGG 5520

TGGTGAAGTT CTTGAGACAT CGTGGCGGAT GTATGGAGCG AGACGATTAG GACAATGGTC 5580

ACCGACGACG GTCACCGCTA TTCAGCACAG AATGGCCCAA CCTGAGTTCT GCTATCAATG 5640

GCCTATTCCG CGTCGCCAGC CCGACTTGCC CCCCAAGCAC GTGTGTCGGG TCGAACCTCG 5700

CTTGCTGGAT GTGGCTTGAC TCTATGGATG TCGCACTCGT AACTCTTTCG CGGTGCGAAG 5760

GGCTTCCCTC TTTCCGCCTG TCCATAGGCC ATTCGCCGTC CCAGCCTTGT CCTCTCGCGT 5820

GCTCCCTCGA AGGTCCCCCT TTGCGGACCA TAGAAATATC AGGACAGCCC AAAGCGGTGG 5880

AGACTGAACT CGCAGCTAAA AACACTACGA GCAGTCCCCC CGCCTCGGAT ACCTTTTTGC 5940

GGTCGTTGCG CCGGAAAAAT GCCAAGGACC GGAAAACGAC CGGAAAACGA GTGTACAAGA 6000

AAGGACGCAA TAGGGGACTA AGACACCTAT TGGCATAATG GCGGAAACTC ACTCGACTAT 6060

GGCGAGCGGC GTCGGCTTGC TGGCTCGCGT CGCTCAGTCA CTCGCTCCTT CGCCTTCTCG 6120

CGGGTTATGC GTTTGGCGGA GAGGGGCGCG CAACCGGCTA AGTAATTACG TCGACCGTGC 6180

TGTCCAAAGG GCTGACCTTT CGCCCGTCAC TCGCGTTGCG TTAATTACAC TCAATCGAGT 6240

GAGTAATCCG TGGGGTCCGA AATGTGAAAT ACGAAGGCCG AGCATACAAC ACACCTTAAC 6300

ACTCGCCTAT TGTTAAAGTG TGTCCTTTGT CGATACTGGT ACTAATGCGG 6350

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 713 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AAGCTTGGGC TGCAGGTCGA TCGACTCTAG AGGATCGATC CCCACCATGG GTCAATCACG 60

CTACCTCCTC TTTTTGGCCA CCCTTGCCCT CCTAAACCAC CTCAGTTTGG CCAGGGTCAT 120

TCCAGTCTCT GGACCTGCCA GGTGTCTTAG CCAGTCCCGA AACCTGCTGA AGACCACAGA 180

TGACATGGTG AAGACGGCCA GAGAAAAACT GAAACATTAT TCCTGCACTG CTGAAGACAT 240

CGATCATGAA GACATCACAC GGGACCAAAC CAGCACATTG AAGACCTGTT TACCACTGGA 300

ACTACACAAG AACGAGAGTT GCCTGGCTAC TAGAGAGACT TCTTCCACAA CAAGAGGGAG 360

CTGCCTGCCC CCACAGAAGA CGTCTTTGAT GATGACCCTG TGCCTTGGTA GCATCTATGA 420

GGACTTGAAG ATGTACCAGA CAGAGTTCCA GGCCATCAAC GCAGCACTTC AGAATCACAA 480

CCATCAGCAG ATCATTCTAG ACAAGGGCAT GCTGGTGGCC ATCGATGAGC TGATGCAGTC 540

TCTGAATCAT AATGGCGAGA CTCTGCGCCA GAAACCTCCT GTGGGAGAAG CAGACCCTTA 600

CAGAGTGAAA ATGAAGCTCT GCATCCTGCT TCACGCCTTC AGCACCCGCG TCGTGACCAT 660

CAACAGGGTG ATGGGCTATC TGAGCTCCGC CTGAGAATTC ATTGATCCAC TAG 713

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 713 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTCGAACCCG ACGTCCAGCT AGCTGAGATC TCCTAGCTAG GGGTGGTACC CAGTTAGTGC 60

GATGGAGGAG AAAAACCGGT GGGAACGGGA GGATTTGGTG GAGTCAAACC GGTCCCAGTA 120

AGGTCAGAGA CCTGGACGGT CCACAGAATC GGTCAGGGCT TTGGACGACT TCTGGTGTCT 180

ACTGTACCAC TTCTGCCGGT CTCTTTTTGA CTTTGTAATA AGGACGTGAC GACTTCTGTA 240

GCTAGTACTT CTGTAGTGTG CCCTGGTTTG GTCGTGTAAC TTCTGGACAA ATGGTGACCT 300

TGATGTGTTC TTGCTCTCAA CGGACCGATG ATCTCTCTGA AGAAGGTGTT GTTCTCCCTC 360

GACGGACGGG GGTGTCTTCT GCAGAAACTA CTACTGGGAC ACGGAACCAT CGTAGATACT 420

CCTGAACTTC TACATGGTCT GTCTCAAGGT CCGGTAGTTG CGTCGTGAAG TCTTAGTGTT 480

GGTAGTCGTC TAGTAAGATC TGTTCCCGTA CGACCACCGG TAGCTACTCG ACTACGTCAG 540

AGACTTAGTA TTACCGCTCT GAGACGCGGT CTTTGGAGGA CACCCTCTTC GTCTGGGAAT 600

GTCTCACTTT TACTTCGAGA CGTAGGACGA AGTGCGGAAG TCGTGGGCGC AGCACTGGTA 660

GTTGTCCCAC TACCCGATAG ACTCGAGGCG GACTCTTAAG TAACTAGGTG ATC 713

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 215 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Gly Gln Ser Arg Tyr Leu Leu Phe Leu Ala Thr Leu Ala Leu Leu
1 5 10 15

Asn His Leu Ser Leu Ala Arg Val Ile Pro Val Ser Gly Pro Ala Arg
20 25 30

Cys Leu Ser Gln Ser Arg Asn Leu Leu Lys Thr Thr Asp Asp Met Val
35 40 45

Lys Thr Ala Arg Glu Lys Leu Lys His Tyr Ser Cys Thr Ala Glu Asp
50 55 60

Ile Asp His Glu Asp Ile Thr Arg Asp Gln Thr Ser Thr Leu Lys Thr
65 70 75 80

Cys Leu Pro Leu Glu Leu His Lys Asn Glu Ser Cys Leu Ala Thr Arg
85 90 95

Glu Thr Ser Ser Thr Thr Arg Gly Ser Cys Leu Pro Pro Gln Lys Thr
100 105 110

Ser Leu Met Met Thr Leu Cys Leu Gly Ser Ile Tyr Glu Asp Leu Lys
115 120 125

Met Tyr Gln Thr Glu Phe Gln Ala Ile Asn Ala Ala Leu Gln Asn His
130 135 140

Asn His Gln Gln Ile Ile Leu Asp Lys Gly Met Leu Val Ala Ile Asp
145 150 155 160

Glu Leu Met Gln Ser Leu Asn His Asn Gly Glu Thr Leu Arg Gln Lys
165 170 175

Pro Pro Val Gly Glu Ala Asp Pro Tyr Arg Val Lys Met Lys Leu Cys
180 185 190

Ile Leu Leu His Ala Phe Ser Thr Arg Val Val Thr Ile Asn Arg Val
195 200 205

Met Gly Tyr Leu Ser Ser Ala
210 215

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1061 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AAGCTTGGGC TGCAGGTCGA TCGACTCTAG AGGATCGATC CCCACCATGG GTCCTCAGAA 60

GCTAACCATC TCCTGGTTTG CCATCGTTTT GCTGGTGTCT CCACTCATGG CCATGTGGGA 120

GCTGGAGAAA GACGTTTATG TTGTAGAGGT GGACTGGACT CCCGATGCCC CTGGAGAAAC 180

AGTGAACCTC ACCTGTGACA CGCCTGAAGA AGATGACATC ACCTGGACCT CAGACCAGAG 240

ACATGGAGTC ATAGGCTCTG GAAAGACCCT GACCATCACT GTCAAAGAGT TTCTAGATGC 300

TGGCCAGTAC ACCTGCCACA AAGGAGGCGA GACTCTGAGC CACTCACATC TGCTGCTCCA 360

CAAGAAGGAA AATGGAATTT GGTCCACTGA AATTTTAAAA AATTTCAAAA ACAAGACTTT 420

CCTGAAGTGT GAAGCACCAA ATTACTCCGG ACGGTTCACG TGCTCATGGC TGGTGCAAAG 480

AAACATGGAC TTGAAGTTCA ACATCAAGAG CAGTAGCAGT TCCCCTGACT CTCGGGCAGT 540

GACATGTGGA ATGGCGTCTC TGTCTGCAGA GAAGGTCACA CTGGACCAAA GGGACTATGA 600

GAAGTATTCA GTGTCCTGCC AGGAGGATGT CACCTGCCCA ACTGCCGAGG AGACCCTGCC 660

CATTGAACTG GCGTTGGAAG CACGGCAGCA GAATAAATAT GAGAACTACA GCACCAGCTT 720

CTTCATCAGG GACATCATCA AACCAGACCC GCCCAAGAAC TTGCAGATGA AGCCTTTGAA 780

GAACTCACAG GTGGAGGTCA GCTGGGAGTA CCCTGACTCC TGGAGCACTC CCCATTCCTA 840

CTTCTCCCTC AAGTTCTTTG TTCGAATCCA GCGCAAGAAA GAAAAGATGA AGGAGACAGA 900

GGAGGGGTGT AACCAGAAAG GTGCGTTCCT CGTAGAGAAG ACATCTACCG AAGTCCAATG 960

CAAAGGCGGG AATGTCTGCG TGCAAGCTCA GGATCGCTAT TACAATTCCT CATGCAGCAA 1020

GTGGGCATGT GTTCCCTGCA GGGTCCGATC CTAGGAATTC C 1061

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1060 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTCGAACCCG ACGTCCAGCT AGCTGAGATC TCCTAGCTAG GGGTGGTACC CAGGAGTCTT 60

CGATTGGTAG AGGACCAAAC GGTAGCAAAA CGACCACAGA GGTGAGTACC GGTACACCCT 120

CGACCTCTTT CTGCAAATAC AACATCTCCA CCTGACCTGA GGGCTACGGG GACCTCTTTG 180

TCACTTGGAG TGGACACTGT GCGGACTTCT TCTACTGTAG TGGACCTGGA GTCTGGTCTC 240

TGTACCTCAG TATCCGAGAC CTTTCTGGGA CTGGTAGTGA CAGTTTCTCA AAGATCTACG 300

ACCGGTCATG TGGACGGTGT TTCCTCCGCT CTGAGACTCG GTGAGTGTAG ACGACGAGGT 360

GTTCTTCCTT TTACCTTAAA CCAGGTGACT TTAAAATTTT TTAAAGTTTT TGTTCTGAAA 420

GGACTTCACA CTTCGTGGTT TAATGAGGCC TGCCAAGTGC ACGAGTACCG ACCACGTTTC 480

TTTGTACCTG AACTTCAAGT TGTAGTTCTC GTCATCGTCA AGGGGACTGA GAGCCCGTCA 540

CTGTACACCT TACCGCAGAG ACAGACGTCT CTTCCAGTGT GACCTGGTTT CCCTGATACT 600

CTTCATAAGT CACAGGACGG TCCTCCTACA GTGGACGGGT TGACGGCTCC TCTGGGACGG 660

GTAACTTGAC CGCAACCTTC GTGCCGTCGT CTTATTTATA CTCTTGATGT CGTGGTCGAA 720

GAAGTAGTCC CTGTAGTAGT TTGGTCTGGG CGGGTTCTTG AACGTCTACT TCGGAAACTT 780

CTTGAGTGTC CACCTCCAGT CGACCCTCAT GGGACTGAGG ACCTCGTGAG GGGTAAGGAT 840

GAAGAGGGAG TTCAAGAAAC AAGCTTAGGT CGCGTTCTTT CTTTTCTACT TCCTCTGTCT 900

CCTCCCCACA TTGGTCTTTC CACGCAAGGA GCATCTCTTC TGTAGATGGC TTCAGGTTAC 960

GTTTCCGCCC TTACAGACGC ACGTTCGAGT CCTAGCGATA ATGTTAAGGA GTACGTCGTT 1020

CACCCGTACA CAAGGGACGT CCCAGGCTAG GATCTTAAGG 1060

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 335 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Gly Pro Gln Lys Leu Thr Ile Ser Trp Phe Ala Ile Val Leu Leu
1 5 10 15

Val Ser Pro Leu Met Ala Met Trp Glu Leu Glu Lys Asp Val Tyr Val
20 25 30

Val Glu Val Asp Trp Thr Pro Asp Ala Pro Gly Glu Thr Val Asn Leu
35 40 45

Thr Cys Asp Thr Pro Glu Glu Asp Asp Ile Thr Trp Thr Ser Asp Gln
50 55 60

Arg His Gly Val Ile Gly Ser Gly Lys Thr Leu Thr Ile Thr Val Lys
65 70 75 80

Glu Phe Leu Asp Ala Gly Gln Tyr Thr Cys His Lys Gly Gly Glu Thr
85 90 95

Leu Ser His Ser His Leu Leu Leu His Lys Lys Glu Asn Gly Ile Trp
100 105 110

Ser Thr Glu Ile Leu Lys Asn Phe Lys Asn Lys Thr Phe Leu Lys Cys
115 120 125

Glu Ala Pro Asn Tyr Ser Gly Arg Phe Thr Cys Ser Trp Leu Val Gln
130 135 140

Arg Asn Met Asp Leu Lys Phe Asn Ile Lys Ser Ser Ser Ser Ser Pro
145 150 155 160

Asp Ser Arg Ala Val Thr Cys Gly Met Ala Ser Leu Ser Ala Glu Lys
165 170 175

```
Val Thr Leu Asp Gln Arg Asp Tyr Glu Lys Tyr Ser Val Ser Cys Gln
            180                 185                 190

Glu Asp Val Thr Cys Pro Thr Ala Glu Glu Thr Leu Pro Ile Glu Leu
        195                 200                 205

Ala Leu Glu Ala Arg Gln Gln Asn Lys Tyr Glu Asn Tyr Ser Thr Ser
    210                 215                 220

Phe Phe Ile Arg Asp Ile Ile Lys Pro Asp Pro Pro Lys Asn Leu Gln
225                 230                 235                 240

Met Lys Pro Leu Lys Asn Ser Gln Val Glu Val Ser Trp Glu Tyr Pro
                245                 250                 255

Asp Ser Trp Ser Thr Pro His Ser Tyr Phe Ser Leu Lys Phe Phe Val
            260                 265                 270

Arg Ile Gln Arg Lys Lys Glu Lys Met Lys Glu Thr Glu Glu Gly Cys
        275                 280                 285

Asn Gln Lys Gly Ala Phe Leu Val Glu Lys Thr Ser Thr Glu Val Gln
    290                 295                 300

Cys Lys Gly Gly Asn Val Cys Val Gln Ala Gln Asp Arg Tyr Tyr Asn
305                 310                 315                 320

Ser Ser Cys Ser Lys Trp Ala Cys Val Pro Cys Arg Val Arg Ser
                325                 330                 335
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 61 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCGCCGGTGG CGGTGGCTCG GGCGGTGGTG GGTCGGGTGG CGGCGGATCT TCCATGGAGC 60

T 61

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCATGGAAGA TCCGCCGCCA CCCGACCCAC CACCGCCCGA GCCACCGCCA CCGGCGGAGC 60

T 61

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CAGAGTGAAA ATGAAGCT 18

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GAAGCTCTGC ATCCTGCT 18

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGGTCCGATC CGGTGGCGGT GGCTCGGGCG GTGGTGGGTC GGGTGGCGGC GGATCTTCCA 60

TG 62

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GATCCATGGA AGATCCGCCG CCACCCGACC CACCACCGCC CGAGCCACCG CCACCGGATC 60

GGACCCTGCA 70

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CTATTACAAT TCCTCATG 18

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GAGGGCAAGG GTGGCCAA 18

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TTGCTGGAGC TCCGCCGGTG GCGGTGGCTC GGGCGGTGGT GGGTCGGGTG GCGGCGGATC 60

TATGTGG 67

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CACATAGATC CGCCGCCACC CGACCCACCA CCGCCCGAGC CACCGCCACC GGCGGAGCTC 60

CAGCAAA 67

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 82 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CTGGCCTGCA GGGTCCGATC CGGTGGCGGT GGCTCGGGCG GTGGTGGGTC GGGTGGCGGC 60

GGATCTAGGG TCATTCCAGT CT 82

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 82 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CTGGAATGAC CCTAGATCCG CCGCCACCCG ACCCACCACC GCCCGAGCCA CCGCCACCGG 60

ATCGGACCCT GCAGGCCAGA GA 82

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GCAAAGGCGG GAATGTCT 18

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AGGAATAATG TTTCAGTT 18

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CAGCAGTGCA GGAATAAT 18

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CCCTGCAGGG TCCGATCCGG TGGCGGTGGC TCGGGCGGTG GTGGGTCGGG TGGCGGCGGA 60

TCTTCCATGG GTCAA 75

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CCCTGCAGGG TCCGATCCGG TGGCGGTGGC TCGGGCGGTG GTGGGTCGGG TGGCGGCGGA 60

TCTTCCATGG GTCAA 75

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TGTGTTCCCT GCAGGGTCCG ATCCGGTGGC GGTGGCTCGG GCGGTGGTGG GTCGGGTGGC 60

GGCGGATCTA GGGTCATTCC AGTCTCTGGA CCTGCC 96

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GTCATCTTCT TCAGGCGT 18

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TGCAGTGGTG GCGGTGGCGG CGGATCTAGA AAC 33

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Gly Gly Gly Gly Gly Gly Ser
1 5

CLAIMS

The invention claimed is:

1. DNA comprising DNA encoding IL-12 p35 subunit, DNA encoding a polypeptide linker and DNA encoding IL-12 p40 subunit, wherein the DNA encoding the polypeptide linker is positioned between the DNA encoding the IL-12 p35 subunit and the DNA encoding the IL-12 p40 subunit and wherein expression of the DNA results in production of a bioactive IL-12 fusion protein comprising the IL-12 p35 subunit and the IL-12 p40 subunit joined by the encoded polypeptide linker.

2. DNA of Claim 1 wherein the IL-12 p35 subunit and the IL-12 p40 subunit are of human or mouse origin and the polypeptide linker is selected from the group consisting of: $(Gly_4Ser)_3$; $(Gly_4Ser)_3Ser$; Gly_6Ser; and $(Gly_4Ser)_2Ser$.

3. DNA encoding a bioactive IL-12 protein, wherein the bioactive IL-12 protein comprises IL-12 p35 subunit and IL-12 p40 subunit joined by a polypeptide linker.

4. DNA of Claim 3 wherein the polypeptide linker is selected from the group consisting of: $(Gly_4Ser)_3$; $(Gly_4Ser)_3Ser$; Gly_6Ser; and $(Gly_4Ser)_2Ser$.

5. DNA encoding a bioactive protein, wherein the bioactive protein comprises two subunits present in a corresponding native dimeric protein and a polypeptide linker, wherein the two subunits are joined in the bioactive protein by a polypeptide linker.

6. DNA of Claim 5 wherein the polypeptide linker is selected from the group consisting of: $(Gly_4Ser)_3$; $(Gly_4Ser)_3Ser$; Gly_6Ser; and $(Gly_4Ser)_2Ser$.

7. A bioactive IL-12 fusion protein encoded by the DNA of Claim 1.

8. A bioactive IL-12 fusion protein encoded by the DNA of Claim 2.

9. A bioactive IL-12 protein encoded by the DNA of Claim 3.

10. A bioactive IL-12 protein which comprises IL-12 p35 subunit and IL-12 p40 subunit joined by a polypeptide linker.

11. A bioactive IL-12 protein of Claim 10 wherein the IL-12 p35 subunit and the IL-12 p40 subunit are of human or mouse origin and the polypeptide linker is 7 to 16 amino acid residues.

12. A bioactive IL-12 protein of Claim 11 wherein the polypeptide linker is selected from the group consisting of: $(Gly_4Ser)_3$; $(Gly_4Ser)_3Ser$; Gly_6Ser; and $(Gly_4Ser)_2Ser$.

13. An expression vector comprising DNA of Claim 1.

14. An expression vector of Claim 13 which is a retrovirus vector.

15. An expression vector of Claim 14 which is an SFG vector.

16. An expression vector of Claim 15 selected from the group consisting of:
    a) pSFG.IL-12.p35.linker.p40;
    b) pSFG.IL-12.p40.linker.p35;
    c) pSFG.IL-12.p35.linker.Δp40;
    d) pSFG.IL-12.p40.linker.Δp35; and
    e) pSFG.hIL-12.p40.linker.Δp35.

17. A method of producing a bioactive IL-12 protein comprising the steps of:
    a) providing an expression vector comprising DNA encoding IL-12 p35 subunit, DNA encoding a polypeptide linker and DNA encoding Il-12 p40 subunit, wherein the DNA encoding the polypeptide linker is positioned between the DNA encoding the IL-12 p35 subunit and the DNA encoding the IL-12 p40 subunit;
    b) introducing the expression vector into an appropriate host cell;
    c) maintaining the host cell resulting from step b) under conditions appropriate for expression of the DNA present in the expression vector, resulting in production of a bioactive IL-12 protein in which the two subunits are joined by the polypeptide linker.

18. The method of Claim 17, wherein the IL-12 p35 subunit and the IL-12 p40 subunit are of human or mouse origin and the polypeptide linker is 7 to 16 amino acid residues.

19. The method of Claim 18, wherein the polypeptide linker is selected from the group consisting of: $(Gly_4Ser)_3$; $(Gly_4Ser)_3Ser$; Gly_6Ser; and $(Gly_4Ser)_2Ser$.

20. The method of Claim 17, wherein the expression vector is a retrovirus vector.

21. The method of Claim 20, wherein the expression vector is an SFG vector.

22. The method of Claim 21, wherein the SFG vector is selected from the group consisting of:
a) pSFG.IL-12.p35.linker.p40;
b) pSFG.IL-12.p40.linker.p35;
c) pSFG.IL-12.p35.linker.Δp40;
d) pSFG.IL-12.p40.linker.Δp35; and
e) pSFG.hIL-12.p40.linker.Δp35.

23. A method of treating a disorder characterized by an established tumor, comprising administering a therapeutically effective dose of IL-12-secreting tumor cells to a subject having a disorder characterized by an established tumor.

24. A method according to Claim 23, wherein the treatment results in reduction of the size of the tumor, prolonged survival of the subject compared with an untreated subject or both.

25. A method according to Claim 23, wherein the IL-12-secreting tumor cells are selected from the group consisting of CMS-5 tumor cells and B16 tumor cells.

26. A method of treating a disorder characterized by an established tumor, comprising administering a therapeutically effective dose of tumor cells secreting a bioactive IL-12 protein which comprises IL-12 p35 subunit and IL-12 p40 subunit joined by a

polypeptide linker to a subject having a disorder characterized by an established tumor.

27. A method of reducing the size of an established tumor in a subject comprising administering a therapeutically effective dose of IL-12-secreting tumor cells to a subject having an established tumor, thereby reducing the size of the established tumor.

28. A method according to Claim 27, wherein the size of the tumor is reduced by greater than 50%.

29. A method according to Claim 27, wherein the established tumor is a melanoma, a fibrosarcoma or a renal cell carcinoma.

30. A method according to Claim 27, wherein the IL-12-secreting tumor cells secrete a bioactive IL-12 protein which comprises IL-12 p35 subunit and IL-12 p40 subunit joined by a polypeptide linker.

31. A method according to Claim 27, wherein the IL-12-secreting tumor cells are of the same type as the established tumor.

32. A method of preventing the establishment of a tumor in a subject, comprising administering a therapeutically effective dose of tumor cells which secrete a bioactive IL-12 protein which comprises IL-12 p35 subunit and IL-12 p40 subunit joined by a polypeptide linker to a subject after the initiation of the tumor and before the establishment of the tumor.

33. Use of IL-12 fusion protein-secreting tumor cells for treating an established tumor.

34. The use of Claim 33, wherein the tumor size is reduced.

35. Use, for the manufacture of a medicament for treating an established tumor, of IL-12-secreting tumor cells.

36. The use of Claim 33, wherein the IL-12 fusion protein-secreting cells are CMS-5 cells, B16 cells or renal cell carcinoma cells.

37. The use of Claim 35, wherein the IL-12 fusion protein-secreting cells are CMS-5 cells, B16 cells or renal cell carcinoma cells.

38. The use of Claim 33, wherein the established tumor is a fibrosarcoma, a melanoma or a renal cell carcinoma.

39. The use of Claim 33, wherein the IL fusion protein-secreting tumor cells secrete a bioactive IL-12 protein which comprises IL-12 p35 subunit and IL-12 p40 subunit joined by a polypeptide linker.

40. A method of treating an established tumor in the subject, comprising administering a therapeutically effective dose of bioactive IL-12 protein which comprises IL-12 p35 subunit and IL-12 p40 subunit joined by a polypeptide linker.

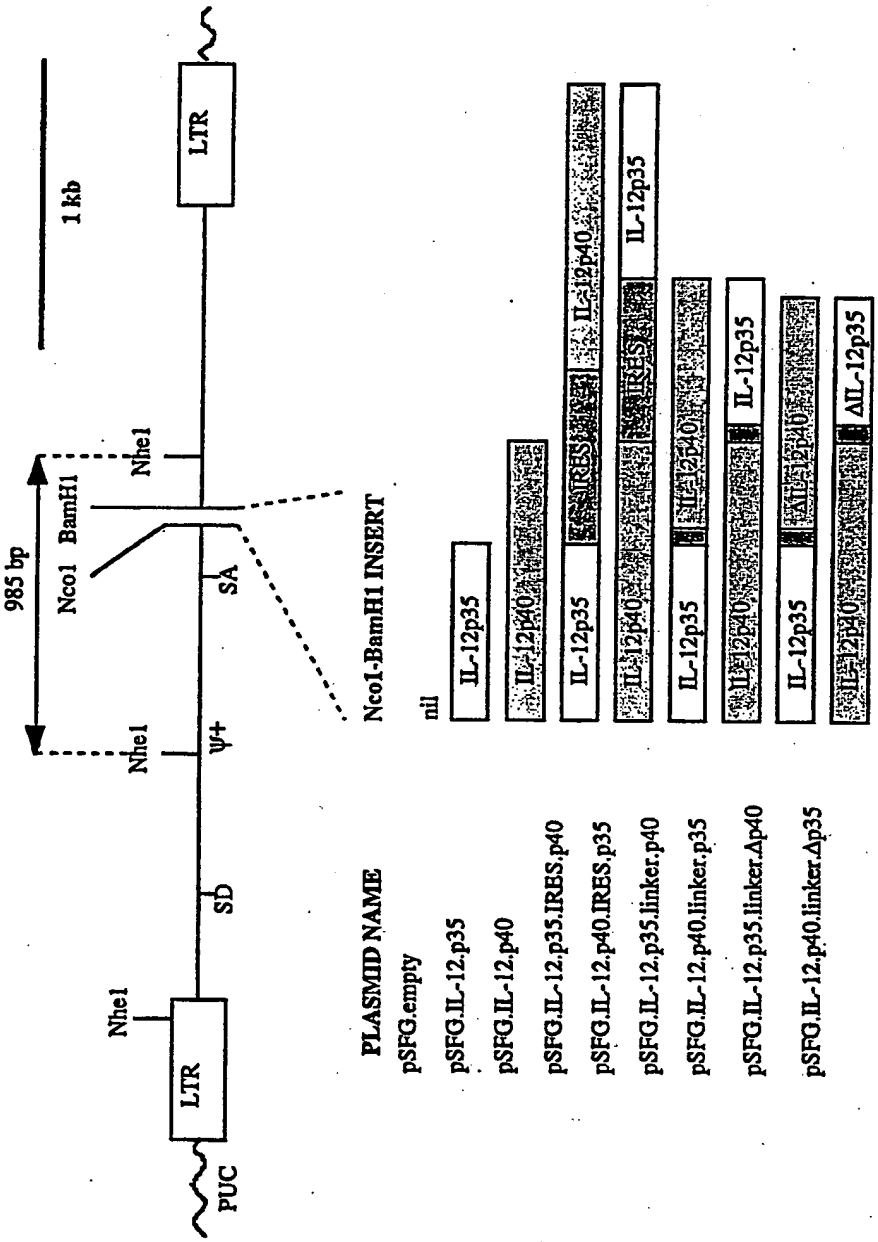

1 kb
LTR
Nhe1
985 bp
BamH1
Nco1
SA
Nhe1
ψ+
SD
Nhe1
LTR
PUC
Nco1-BamH1 INSERT
PLASMID NAME
pSFG.empty
pSFG.IL-12.p35
pSFG.IL-12.p40
pSFG.IL-12.p35.IRES.p40
pSFG.IL-12.p40.IRES.p35
pSFG.IL-12.p35.linker.p40
pSFG.IL-12.p40.linker.p35
pSFG.IL-12.p35.linker.Δp40
pSFG.IL-12.p40.linker.Δp35
nil
IL-12p35
IL-12p40
IRES
ΔIL-12p35

FIGURE 1

```
   IL-12p35          Linker
A  5'>>>AGC.TCC.GCC.GGT.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.TCC.ATG.GGT.CCT.CAG>>>-3'
                                                                IL-12p40
                   Gly.Gly.Gly.Gly.Ser.Gly.Gly.Gly.Gly.Ser.Ser

   IL-12p40          Linker
B  5'>>>GTC.CGA.TCC.GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.TCC.ATG.GGT.CAA>>>-3'
                                                                IL-12p35
                   Gly.Gly.Gly.Gly.Ser.Gly.Gly.Gly.Gly.Ser.Gly.Gly.Gly.Gly.Ser.Ser

   IL-12p35          Linker
C  5'>>>AGC.TCC.GCC.GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.ATG.TGG.GAG.CTG>>>-3'
                                                                ΔIL-12p40
                   Gly.Gly.Gly.Gly.Ser.Gly.Gly.Gly.Gly.Ser.Gly.Gly.Gly.Gly.Ser

   IL-12p40          Linker
D  5'>>>GTC.CGA.TCC.GGT.GGC.GGT.GGC.TCG.GGC.GGT.GGT.GGG.TCG.GGT.GGC.GGC.GGA.TCT.AGG.GTC.ATT.CCA>>>-3'
                                                                ΔIL-12p35
                   Gly.Gly.Gly.Gly.Ser.Gly.Gly.Gly.Gly.Ser.Gly.Gly.Gly.Gly.Ser

   IL-12p40      Linker                    ΔIL-12p35
E  5'>>>TGC.AGT.GGT.GGC.GGT.GGC.GGC.GGA.TCT.AGA.AAC>>>3'
                 Gly.Gly.Gly.Gly.Gly.Gly.Ser
```

FIGURE 2

Positions of Restriction Endonucleases sites (unique sites underlined)

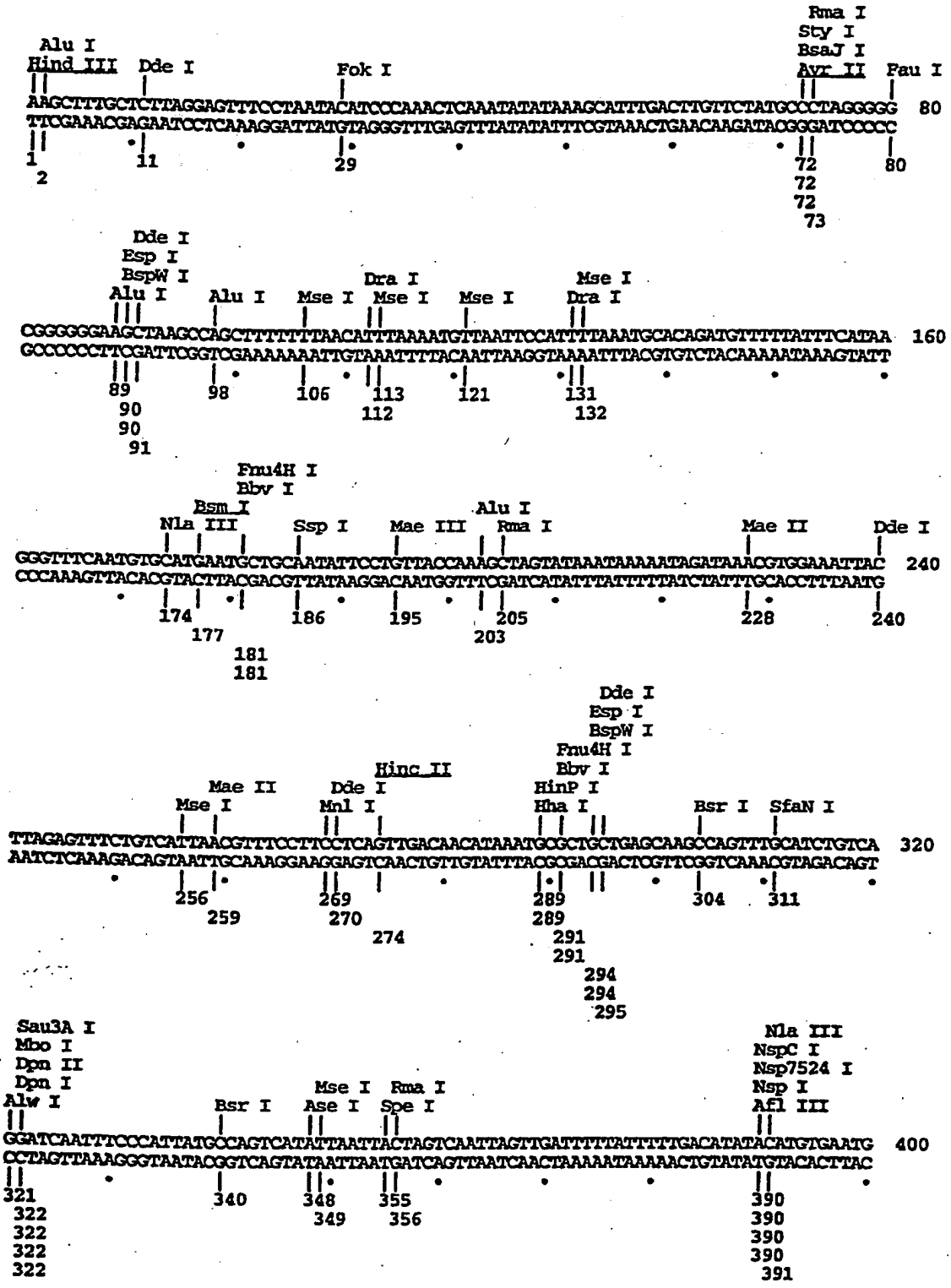


FIGURE 3A

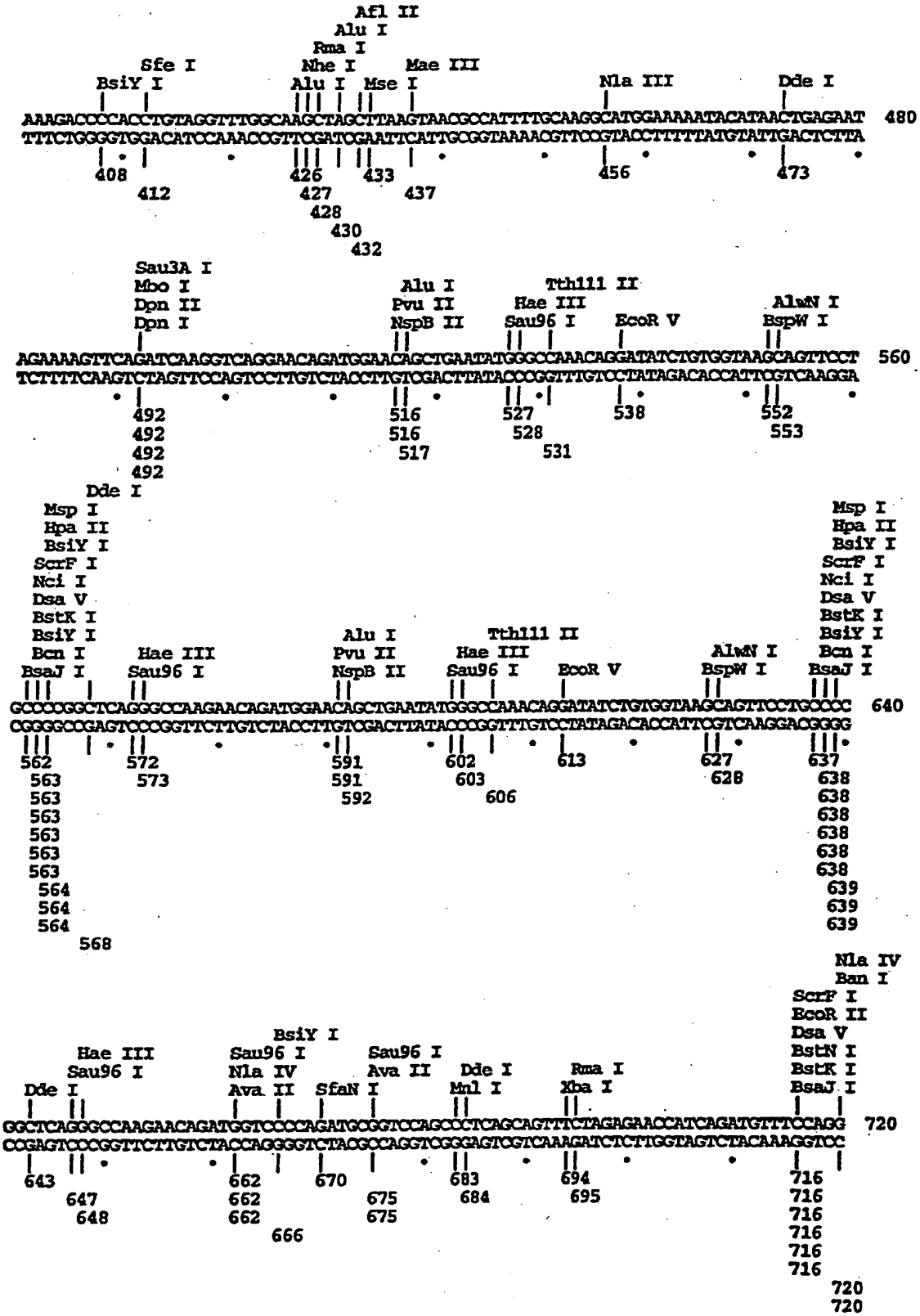

Afl II
Alu I
Rma I
Sfe I
Nhe I
Mae III
BsiY I
Alu I
Mse I
Nla III
Dde I
AAAGACCCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAAGGCATGGAAAAATACATAACTGAGAAT 480
TTTCTGGGGTGGACATCCAAACCGTTCGATCGAATTCATTGCGGTAAAACGTTCCGTACCTTTTTATGTATTGACTCTTA
408
412
426
427
428
430
432
433
437
456
473
Sau3A I
Mbo I
Dpn II
Dpn I
Alu I
Pvu II
NspB II
Tth111 II
Hae III
Sau96 I
EcoR V
AlwN I
BspW I
AGAAAAGTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCT 560
TCTTTTCAAGTCTAGTTCCAGTCCTTGTCTACCTTGTCGACTTATACCCGGTTTGTCCTATAGACACCATTCGTCAAGGA
492
492
492
492
516
516
517
527
528
531
538
552
553
Dde I
Msp I
Hpa II
BsiY I
ScrF I
Nci I
Dsa V
BstK I
BsiY I
Bcn I
BsaJ I
Hae III
Sau96 I
Alu I
Pvu II
NspB II
Tth111 II
Hae III
Sau96 I
EcoR V
AlwN I
BspW I
GCCCCGGCTCAGGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCC 640
CGGGGCCGAGTCCCGGTTCTTGTCTACCTTGTCGACTTATACCCGGTTTGTCCTATAGACACCATTCGTCAAGGACGGGG
562
563
563
563
563
563
563
564
564
564
568
572
573
591
591
592
602
603
606
613
627
628
637
638
638
638
638
638
638
639
639
639
Nla IV
Ban I
ScrF I
EcoR II
Dsa V
BstN I
BstK I
BsaJ I
BsiY I
Hae III
Sau96 I
Dde I
Sau96 I
Nla IV
Ava II
SfaN I
Sau96 I
Ava II
Dde I
Mnl I
Rma I
Xba I
GGCTCAGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGG 720
CCGAGTCCCGGTTCTTGTCTACCAGGGGTCTACGCCAGGTCGGGAGTCGTCAAAGATCTCTTGGTAGTCTACAAAGGTCC
643
647
648
662
662
662
666
670
675
675
683
684
694
695
716
716
716
716
716
716
720
720

FIGURE 3B

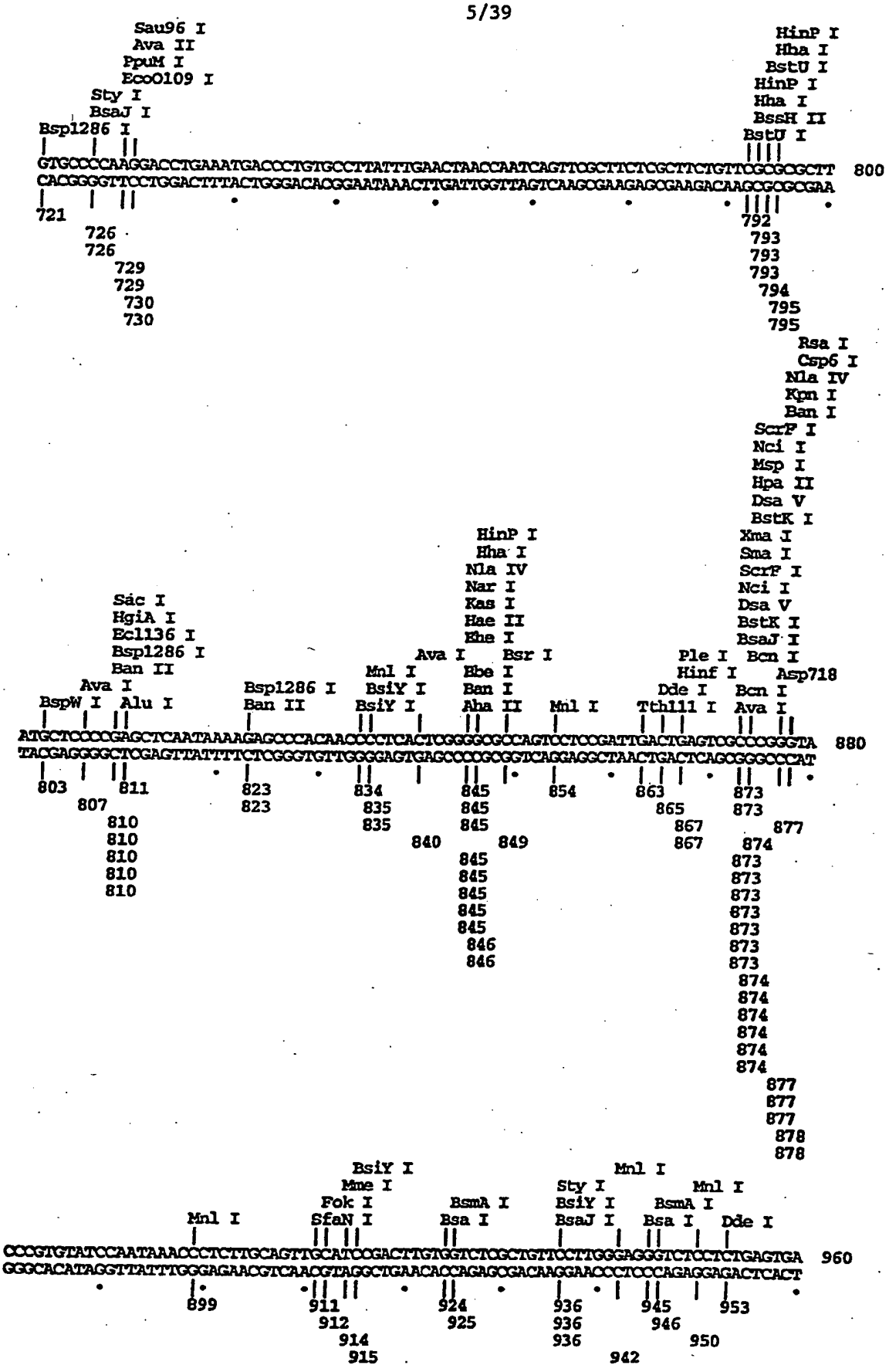

Sau96 I
Ava II
PpuM I
Eco0109 I
Sty I
BsaJ I
Bsp1286 I
HinP I
Hha I
BstU I
HinP I
Hha I
BssH II
BstU I
GTGCCCCAAGGACCTGAAATGACCCTGTGCCTTATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTT 800
CACGGGGTTCCTGGACTTTACTGGGACACGGAATAAACTTGATTGGTTAGTCAAGCGAAGAGCGAAGACAAGCGCGCGAA
721
726
726
729
729
730
730
792
793
793
793
794
795
795
Rsa I
Csp6 I
Nla IV
Kpn I
Ban I
ScrF I
Nci I
Msp I
Hpa II
Dsa V
BstK I
HinP I
Hha I
Nla IV
Nar I
Kas I
Hae II
Ehe I
Xma I
Sma I
ScrF I
Nci I
Dsa V
BstK I
BsaJ I
Sac I
HgiA I
Ecl136 I
Bsp1286 I
Ban II
Ava I
Mnl I
BsiY I
Ava I
Bbe I
Ban I
Aha II
Bsr I
Mnl I
Ple I
Hinf I
Bcn I
Asp718
Dde I
Bcn I
Ava I
Bsp1286 I
Ban II
BsiY I
BspW I
Alu I
ATGCTCCCCGAGCTCAATAAAAGAGCCCACAACCCCTCACTCGGGGCGCCAGTCCTCCGATTGACTGAGTCGCCCGGGTA 880
TACGAGGGGCTCGAGTTATTTTCTCGGGTGTTGGGGAGTGAGCCCCGCGGTCAGGAGGCTAACTGACTCAGCGGGCCCAT
803
807
811
823
823
834
835
835
840
845
845
845
845
845
845
845
845
846
846
849
854
863
865
867
867
873
873
874
873
873
873
873
873
873
873
874
874
874
874
874
874
877
877
877
878
878
877
BsiY I
Mme I
Fok I
SfaN I
Mnl I
BsmA I
Bsa I
Sty I
BsiY I
BsaJ I
Mnl I
BsmA I
Bsa I
Dde I
CCCGTGTATCCAATAAACCCTCTTGCAGTTGCATCCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGA 960
GGGCACATAGGTTATTTGGGAGAACGTCAACGTAGGCTGAACACCAGAGCGACAAGGAACCCTCCCAGAGGAGACTCACT
899
911
912
914
915
924
925
936
936
936
942
945
946
950
953

FIGURE 3C

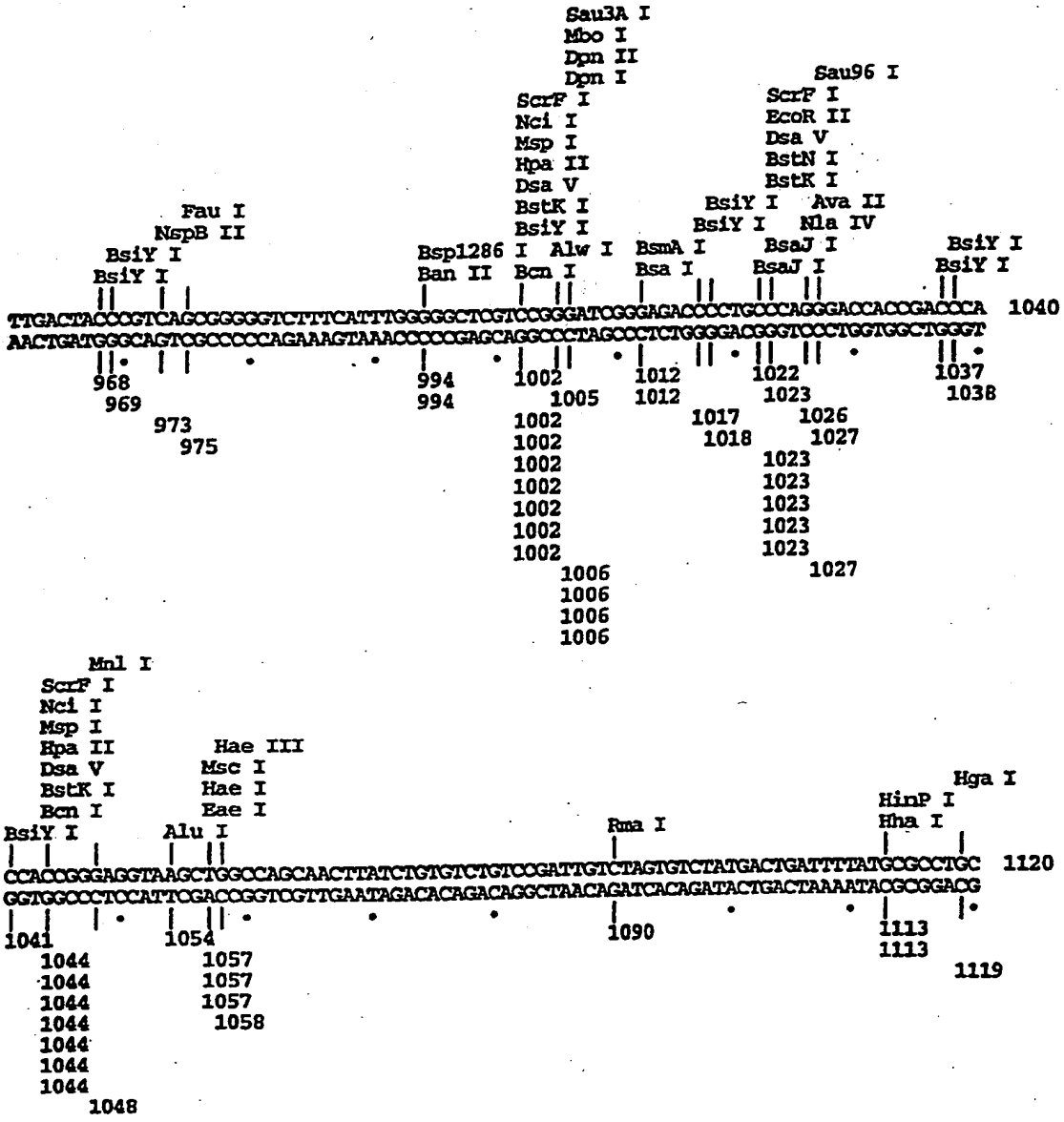

Sau3A I
Mbo I
Dpn II
Dpn I
Sau96 I
ScrF I
Nci I
Msp I
Hpa II
Dsa V
BstK I
BsiY I
ScrF I
EcoR II
Dsa V
BstN I
BstK I
BsiY I Ava II
BsiY I Nla IV
Fau I
NspB II
BsiY I
BsiY I
Bsp1286 I Alw I
Ban II Bcn I
BsmA I
Bsa I
BsaJ I
BsaJ I
BsiY I
BsiY I
TTGACTACCCGTCAGCGGGGGTCTTTCATTTGGGGGCTCGTCCGGGATCGGGAGACCCCTGCCCAGGGACCACCGACCCA 1040
AACTGATGGGCAGTCGCCCCCCAGAAAGTAAACCCCCGAGCAGGCCCTAGCCCTCTGGGGACGGGTCCCTGGTGGCTGGGT
968
969
973
975
994
994
1002
1002
1002
1002
1002
1002
1002
1005
1006
1006
1006
1006
1012
1012
1017
1018
1022
1023
1023
1023
1023
1023
1023
1026
1027
1027
1037
1038
Mnl I
ScrF I
Nci I
Msp I
Hpa II
Dsa V
BstK I
Bcn I
Hae III
Msc I
Hae I
Eae I
BsiY I
Alu I
Rma I
Hga I
HinP I
Hha I
CCACCGGGAGGTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGACTGATTTTATGCGCCTGC 1120
GGTGGCCCTCCATTCGACCGGTCGTTGAATAGACACAGACAGGCTAACAGATCACAGATACTGACTAAAATACGCGGACG
1041
1044
1044
1044
1044
1044
1044
1044
1048
1054
1057
1057
1057
1058
1090
1113
1113
1119

FIGURE 3D

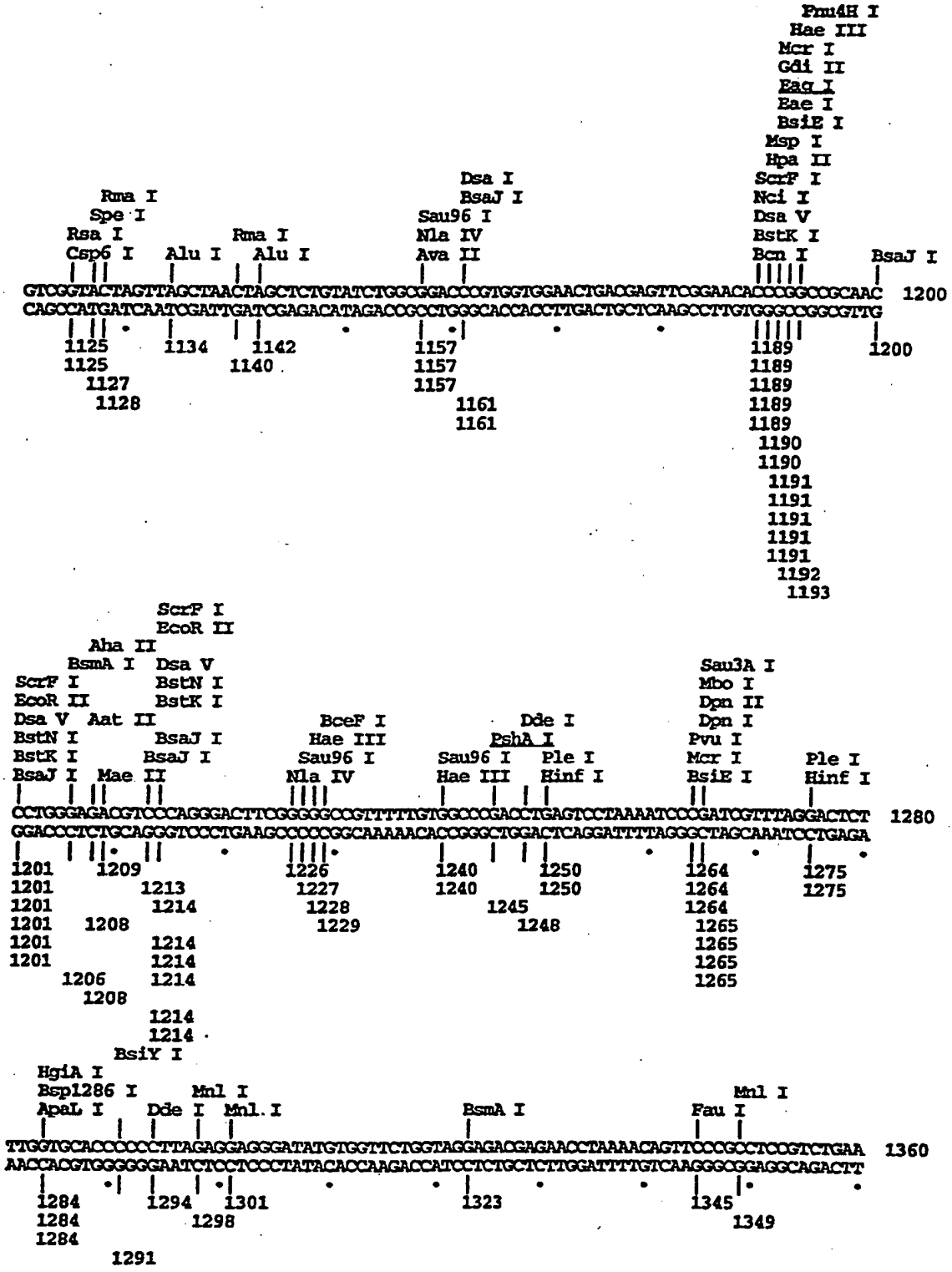

Pmi4H I
Hae III
Mcr I
Gdi II
Eag I
Eae I
BsiE I
Msp I
Hpa II
ScrF I
Nci I
Dsa V
BstK I
Bcn I
BsaJ I
Rma I
Spe I
Rsa I
Csp6 I
Alu I
Rma I
Alu I
Dsa I
BsaJ I
Sau96 I
Nla IV
Ava II
GTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTGGTGGAACTGACGAGTTCGGAACACCCGGCCGCAAC 1200
CAGCCATGATCAATCGATTGATCGAGACATAGACCGCCTGGGCACCACCTTGACTGCTCAAGCCTTGTGGGCCGGCGTTG
1125
1125
1127
1128
1134
1142
1140
1157
1157
1157
1161
1161
1189
1189
1189
1189
1189
1190
1190
1191
1191
1191
1191
1191
1192
1193
1200
ScrF I
EcoR II
Aha II
BsmA I
Dsa V
ScrF I
BstN I
EcoR II
BstK I
Dsa V
Aat II
BstN I
BsaJ I
BstK I
BsaJ I
Mae II
BceF I
Hae III
Sau96 I
Nla IV
Dde I
PshA I
Sau96 I
Hae III
Ple I
Hinf I
Sau3A I
Mbo I
Dpn II
Dpn I
Pvu I
Mcr I
BsiE I
Ple I
Hinf I
CCTGGGAGACGTCCCAGGGACTTCGGGGGCCGTTTTTGTGGCCCGACCTGAGTCCTAAAATCCCGATCGTTTAGGACTCT 1280
GGACCCTCTGCAGGGTCCCTGAAGCCCCCGGCAAAAACACCGGGCTGGACTCAGGATTTTAGGGCTAGCAAATCCTGAGA
1201
1201
1201
1201
1201
1201
1206
1208
1209
1213
1214
1208
1214
1214
1214
1214
1226
1227
1228
1229
1240
1240
1245
1248
1250
1250
1264
1264
1264
1265
1265
1265
1265
1275
1275
BsiY I
HgiA I
Bsp1286 I
ApaL I
Mnl I
Dde I
Mnl I
BsmA I
Mnl I
Fau I
TTGGTGCACCCCCCTTAGAGGAGGGATATGTGGTTCTGGTAGGAGACGAGAACCTAAAACAGTTCCCGCCTCCGTCTGAA 1360
AACCACGTGGGGGGAATCTCCTCCCTATACACCAAGACCATCCTCTGCTCTTGGATTTTGTCAAGGGCGGAGGCAGACTT
1284
1284
1284
1291
1294
1298
1301
1323
1345
1349

FIGURE 3E

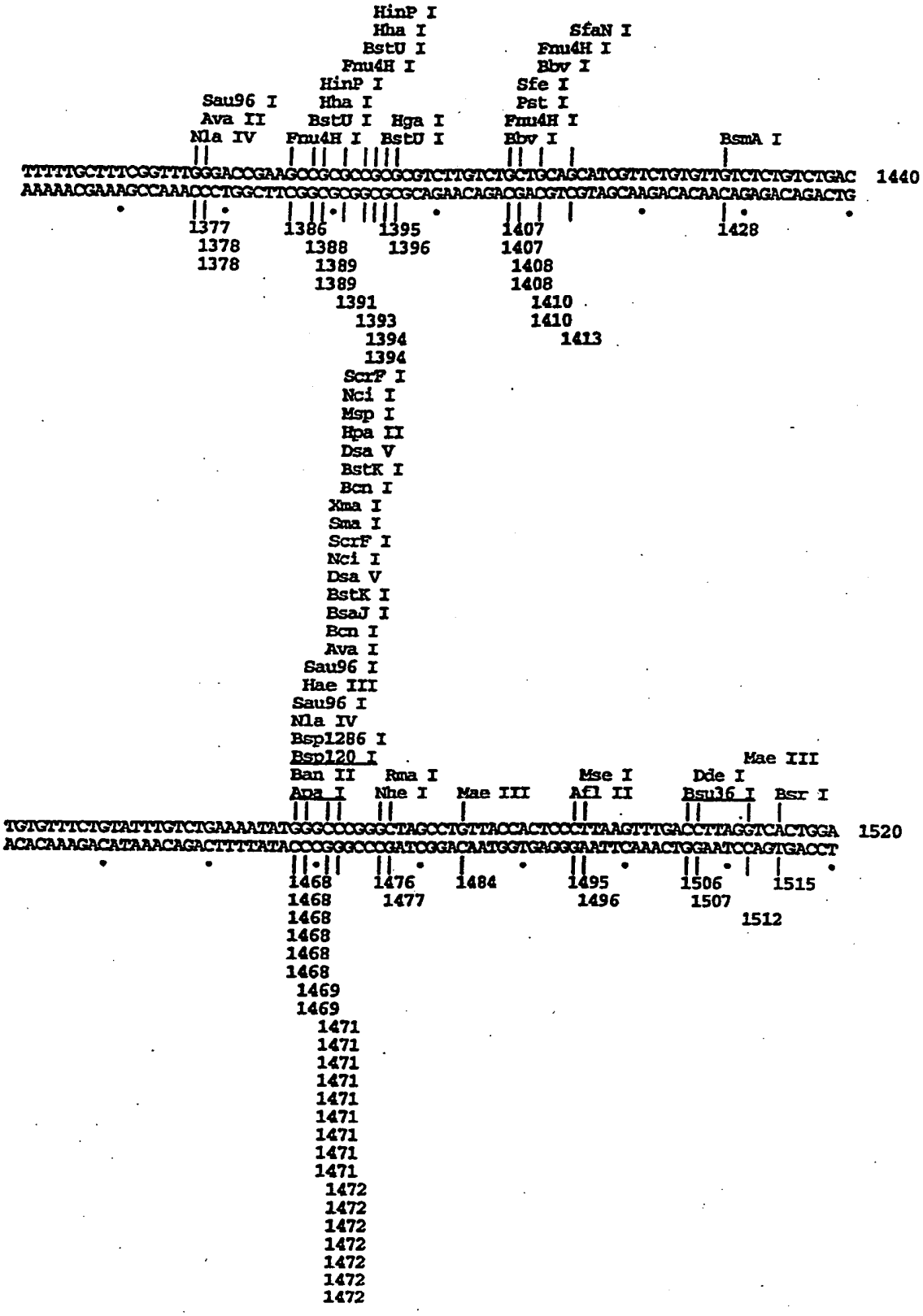

HinP I
Hha I
BstU I
Fnu4H I
HinP I
Hha I
BstU I Hga I
Fnu4H I BstU I
Sau96 I
Ava II
Nla IV
SfaN I
Fnu4H I
Bbv I
Sfe I
Pst I
Fnu4H I
Bbv I
BsmA I
TTTTGCTTTCGGTTTGGGACCGAAGCCGCGCCGCGTCTTGTCTGCTGCAGCATCGTTCTGTGTTGTCTCTGTCTGAC 1440
AAAAACGAAAGCCAAACCCTGGCTTCGGCGCGGCGCAGAACAGACGACGTCGTAGCAAGACACAACAGAGACAGACTG
1377
1378
1378
1386
1388
1389
1389
1391
1393
1394
1394
1395
1396
1407
1407
1408
1408
1410
1410
1413
1428
ScrF I
Nci I
Msp I
Hpa II
Dsa V
BstK I
Bcn I
Xma I
Sma I
ScrF I
Nci I
Dsa V
BstK I
BsaJ I
Bcn I
Ava I
Sau96 I
Hae III
Sau96 I
Nla IV
Bsp1286 I
Bsp120 I
Ban II
Apa I
Rma I
Nhe I
Mae III
Mse I
Afl II
Dde I
Bsu36 I
Mae III
Bsr I
TGTGTTTCTGTATTTGTCTGAAAATATGGGCCCGGGCTAGCCTGTTACCACTCCCTTAAGTTTGACCTTAGGTCACTGGA 1520
ACACAAAGACATAAACAGACTTTTATACCCGGGCCCGATCGGACAATGGTGAGGGAATTCAAACTGGAATCCAGTGACCT
1468
1468
1468
1468
1468
1468
1469
1469
1471
1471
1471
1471
1471
1471
1471
1471
1471
1472
1472
1472
1472
1472
1472
1472
1476
1477
1484
1495
1496
1506
1507
1512
1515

FIGURE 3F

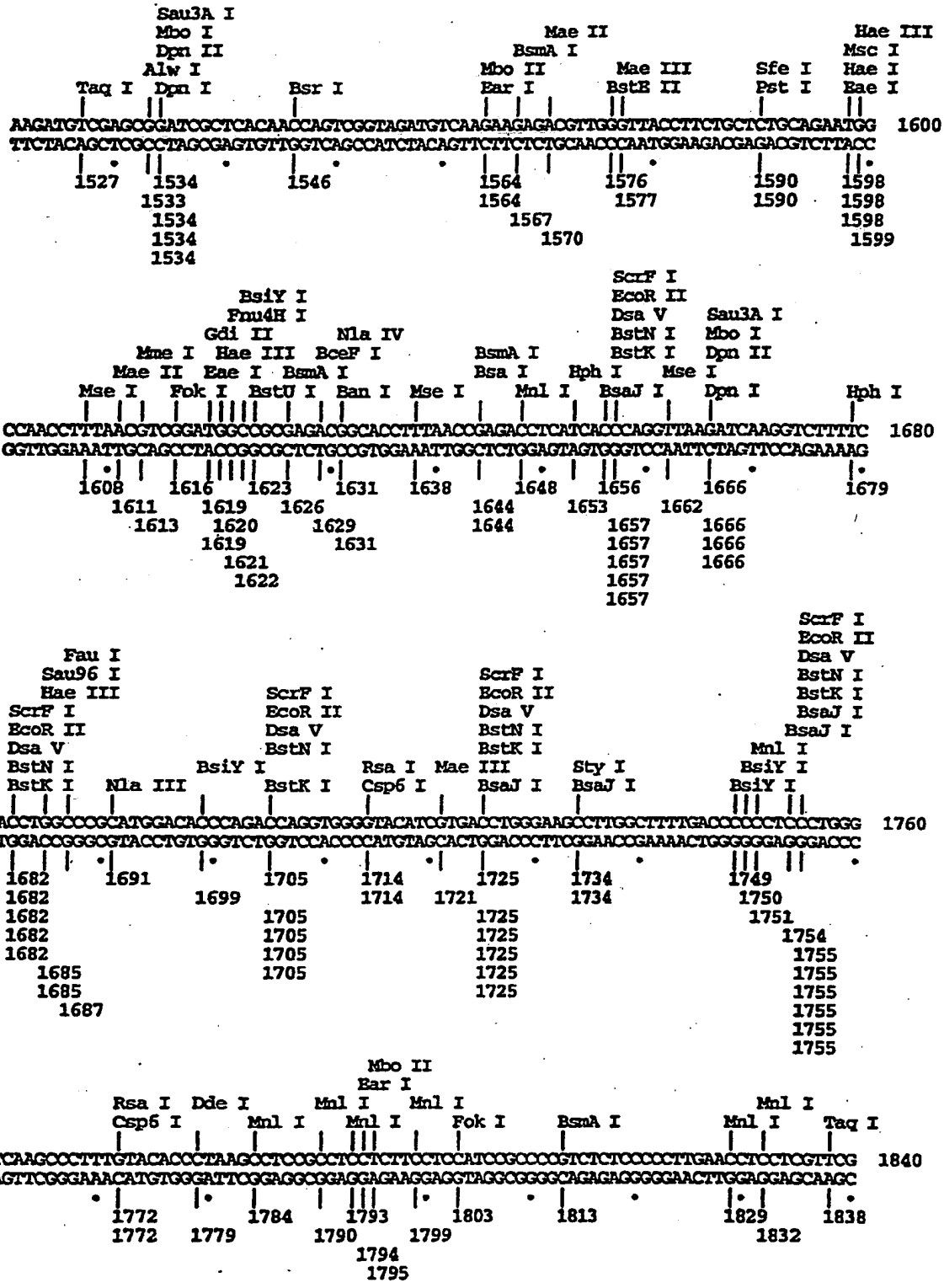

Sau3A I
Mbo I
Dpn II
Alw I
Taq I
Dpn I
Bsr I
Mae II
BsmA I
Mbo II
Bar I
Mae III
BstE II
Sfe I
Pst I
Hae III
Msc I
Hae I
Bae I
AAGATGTCGAGCGGATCGCTCACAACCAGTCGGTAGATGTCAAGAAGAGACGTTGGGTTACCTTCTGCTCTGCAGAATGG 1600
TTCTACAGCTCGCCTAGCGAGTGTTGGTCAGCCATCTACAGTTCTTCTCTGCAACCCAATGGAAGACGAGACGTCTTACC
1527 1534 1546 1564 1576 1590 1598
1533 1564 1577 1590 1598
1534 1567 1598
1534 1570 1599
1534
BsiY I
Pnu4H I
Gdi II
Nla IV
Mme I
Hae III
BceF I
Mae II
Eae I
BsmA I
Mse I
Fok I
BstU I
Ban I
Mse I
BsmA I
Bsa I
Mnl I
Hph I
BsaJ I
ScrF I
EcoR II
Dsa V
BstN I
BstK I
Sau3A I
Mbo I
Dpn II
Mse I
Dpn I
Hph I
CCAACCTTTAACGTGGATGGCGGCGAGACGGCACCTTTAACCGAGACCTCATCACCCAGGTTAAGATCAAGGTCTTTTC 1680
GGTTGGAAATTGCAGCCTACCGGCGCTCTGCCGTGGAAATTGGCTCTGGAGTAGTGGGTCCAATTCTAGTTCCAGAAAAG
1608 1616 1623 1631 1638 1648 1656 1666 1679
1611 1619 1626 1644 1653 1662
1613 1620 1629 1644 1657 1666
1619 1631 1657 1666
1621 1657 1666
1622 1657
1657
ScrF I
EcoR II
Dsa V
BstN I
BstK I
BsaJ I
Fau I
Sau96 I
Hae III
ScrF I
EcoR II
Dsa V
BstN I
BsiY I
Nla III
ScrF I
EcoR II
Dsa V
BstN I
BstK I
Rsa I
Csp6 I
Mae III
BsaJ I
Sty I
BsaJ I
Mnl I
BsiY I
BsiY I
ACCTGGCCCGCATGGACACCCAGACCAGGTGGGGTACATCGTGACCTGGGAAGCCTTGGCTTTTGACCCCCCTCCCTGGG 1760
TGGACCGGGCGTACCTGTGGGTCTGGTCCACCCCATGTAGCACTGGACCCTTCGGAACCGAAAACTGGGGGGGAGGGACCC
1682 1691 1705 1714 1725 1734 1749
1682 1699 1714 1721 1734 1750
1682 1705 1725 1751
1682 1705 1725 1754
1682 1705 1725 1755
1685 1725 1755
1685 1755
1687 1755
1755
Mbo II
Bar I
Rsa I
Csp6 I
Dde I
Mnl I
Mnl I
Mnl I
Fok I
BsmA I
Mnl I
Mnl I
Taq I
TCAAGCCCTTTGTACACCCTAAGCCTCCGCCTCCTCTTCCTCCATCCGCCCCGTCTCTCCCCCTTGAACCTCCTCGTTCG 1840
AGTTCGGGAAACATGTGGGATTCGGAGGCGGAGGAGAAGGAGGTAGGCGGGGCAGAGAGGGGGAACTTGGAGGAGCAAGC
1772 1784 1793 1803 1813 1829 1838
1772 1779 1790 1799 1832
1794
1795

FIGURE 3G

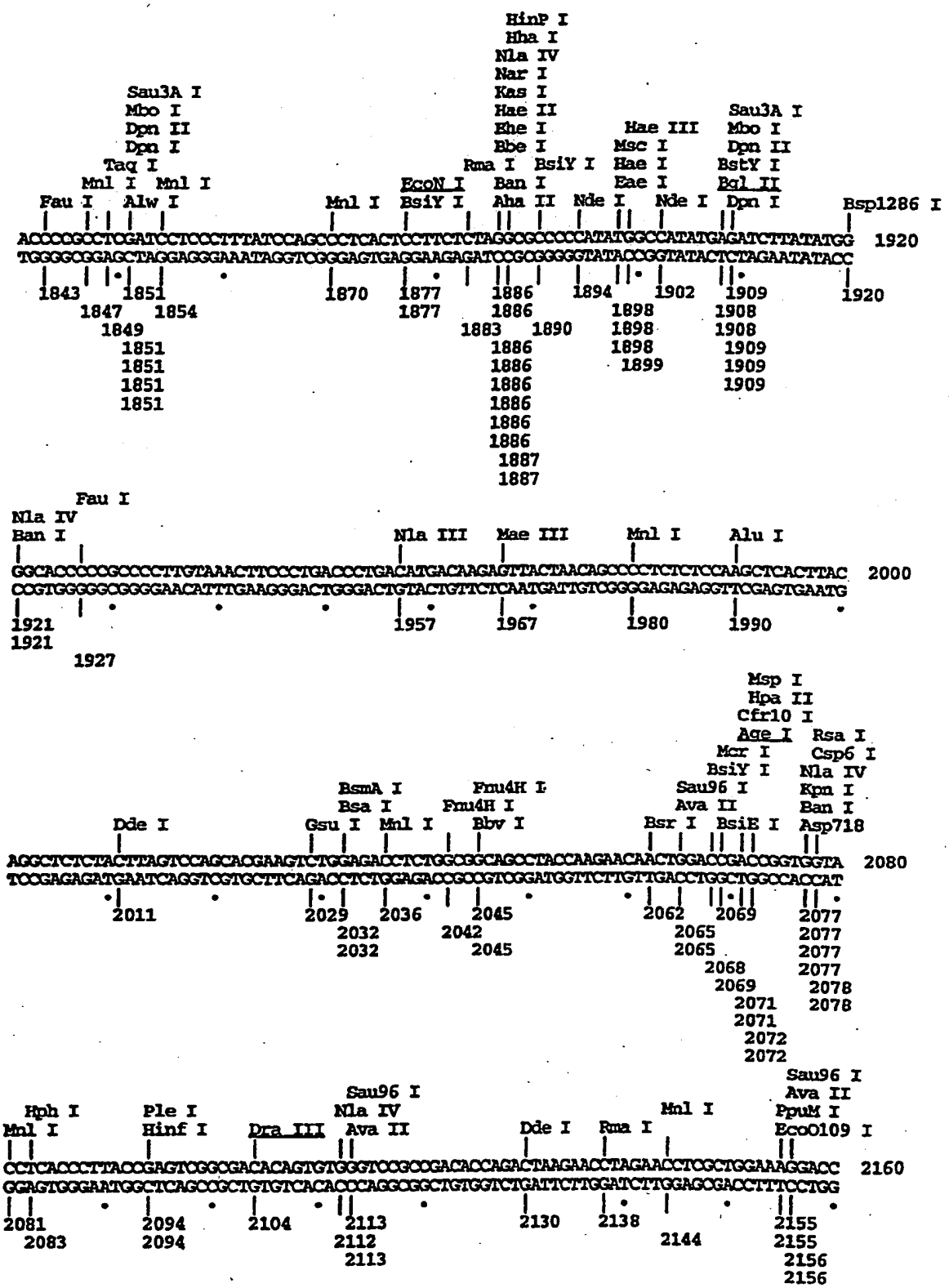

HinP I
Hha I
Nla IV
Nar I
Kas I
Hae II
Ehe I
Bbe I
Rma I BsiY I
EcoN I
Ban I
Aha II
Nde I
Hae III
Msc I
Hae I
Eae I
Nde I
Sau3A I
Mbo I
Dpn II
Dpn I
BstY I
Bgl II
Dpn I
Sau3A I
Mbo I
Dpn II
Dpn I
Taq I
Mnl I Mnl I
Fau I Alw I
Mnl I
Bsp1286 I
ACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCCTTCTCTAGGCGCCCCCATATGGCCATATGAGATCTTATATGG 1920
TGGGGCGGAGCTAGGAGGGAAATAGGTCGGGAGTGAGGAAGAGATCCGCGGGGGTATACCGGTATACTCTAGAATATACC
1843 1851 1870 1877 1877 1886 1886 1883 1890 1894 1898 1898 1898 1899 1902 1909 1908 1908 1909 1909 1909 1920
1847 1854
1849
1851
1851
1851
1851
1886
1886
1886
1886
1886
1886
1887
1887
Fau I
Nla IV
Ban I
Nla III
Mae III
Mnl I
Alu I
GGCACCCCCGCCCCTTGTAAACTTCCCTGACCCTGACATGACAAGAGTTACTAACAGCCCCTCTCTCCAAGCTCACTTAC 2000
CCGTGGGGGCGGGGAACATTTGAAGGGACTGGGACTGTACTGTTCTCAATGATTGTCGGGGAGAGAGGTTCGAGTGAATG
1921 1957 1967 1980 1990
1921
1927
Msp I
Hpa II
Cfr10 I
Age I
Rsa I
Mcr I
Csp6 I
BsiY I
Nla IV
Sau96 I
Kpn I
Ava II
Ban I
BsmA I
Bsa I
Fmu4H I
Fmu4H I
Bsr I
BsiE I
Asp718
Dde I
Gsu I
Mnl I
Bbv I
AGGCTCTCTACTTAGTCCAGCACGAAGTCTGGAGACCTCTGGCGGCAGCCTACCAAGAACAACTGGACCGACCGGTGGTA 2080
TCCGAGAGATGAATCAGGTCGTGCTTCAGACCTCTGGAGACCGCCGTCGGATGGTTCTTGTTGACCTGGCTGGCCACCAT
2011 2029 2036 2045 2062 2069 2077
2032 2042 2065 2077
2032 2045 2065 2077
2068 2077
2069 2078
2071 2078
2071
2072
2072
Sau96 I
Ava II
Sau96 I
Nla IV
Hph I
Mnl I
Ple I
Hinf I
Dra III
Ava II
Dde I
Rma I
Mnl I
PpuM I
Eco0109 I
CCTCACCCTTACCGAGTCGGCGACACAGTGTGGGTCCGCCGACACCAGACTAAGAACCTAGAACCTCGCTGGAAAGGACC 2160
GGAGTGGGAATGGCTCAGCCGCTGTGTCACACCCAGGCGGCTGTGGTCTGATTCTTGGATCTTGGAGCGACCTTTCCTGG
2081 2094 2104 2113 2130 2138 2144 2155 2156
2083 2094 2112 2113 2155 2156

FIGURE 3H

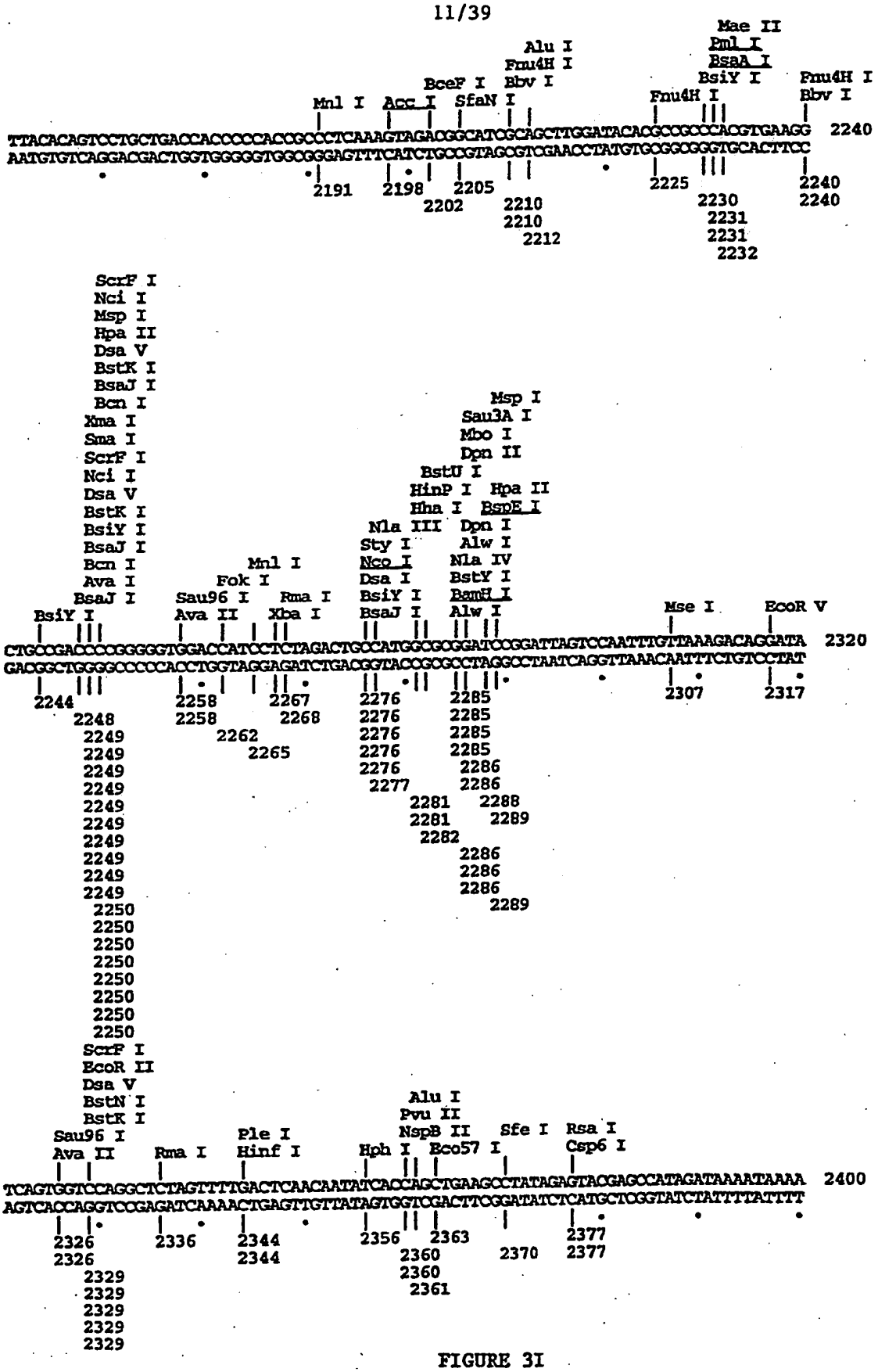

Mae II
Pml I
BsaA I
Alu I
Fmu4H I
Bbv I
BsiY I
Fmu4H I
Bbv I
BceF I
Mnl I
Acc I
SfaN I
Fnu4H I
TTACACAGTCCTGCTGACCACCCCCACCGCCCTCAAAGTAGACGGCATCGCAGCTTGGATACACGCCGCCCACGTGAAGG 2240
AATGTGTCAGGACGACTGGTGGGGGTGGCGGGAGTTTCATCTGCCGTAGCGTCGAACCTATGTGCGGCGGGTGCACTTCC
2191
2198
2205
2202
2210
2210
2212
2225
2230
2231
2231
2232
2240
2240
ScrF I
Nci I
Msp I
Hpa II
Dsa V
BstK I
BsaJ I
Bcn I
Xma I
Sma I
ScrF I
Nci I
Dsa V
BstK I
BsiY I
BsaJ I
Bcn I
Ava I
BsaJ I
Mnl I
Fok I
Sau96 I
Ava II
Rma I
Xba I
BsiY I
Msp I
Sau3A I
Mbo I
Dpn II
BstU I
HinP I
Hpa II
Hha I
BspE I
Nla III
Dpn I
Sty I
Alw I
Nco I
Nla IV
Dsa I
BstY I
BsiY I
BamH I
BsaJ I
Alw I
Mse I
EcoR V
BsiY I
CTGCCGACCCCCGGGGGTGGACCATCCTCTAGACTGCCATGGCGCGGATCCGGATTAGTCCAATTTGTTAAAGACAGGATA 2320
GACGGCTGGGGCCCCCACCTGGTAGGAGATCTGACGGTACCGCGCCTAGGCCTAATCAGGTTAAACAATTTCTGTCCTAT
2244
2248
2249
2249
2249
2249
2249
2249
2249
2249
2249
2250
2250
2250
2250
2250
2250
2250
2250
2258
2258
2262
2265
2267
2268
2276
2276
2276
2276
2276
2277
2281
2281
2282
2285
2285
2285
2285
2286
2286
2286
2286
2289
2288
2289
2307
2317
ScrF I
EcoR II
Dsa V
BstN I
BstK I
Alu I
Pvu II
NspB II
Sfe I
Rsa I
Sau96 I
Ava II
Rma I
Ple I
Hinf I
Hph I
Eco57 I
Csp6 I
TCAGTGGTCCAGGCTCTAGTTTTGACTCAACAATATCACCAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAA 2400
AGTCACCAGGTCCGAGATCAAAACTGAGTTGTTATAGTGGTCGACTTCGGATATCTCATGCTCGGTATCTATTTTATTTT
2326
2326
2329
2329
2329
2329
2329
2336
2344
2344
2356
2360
2360
2361
2363
2370
2377
2377

FIGURE 3I

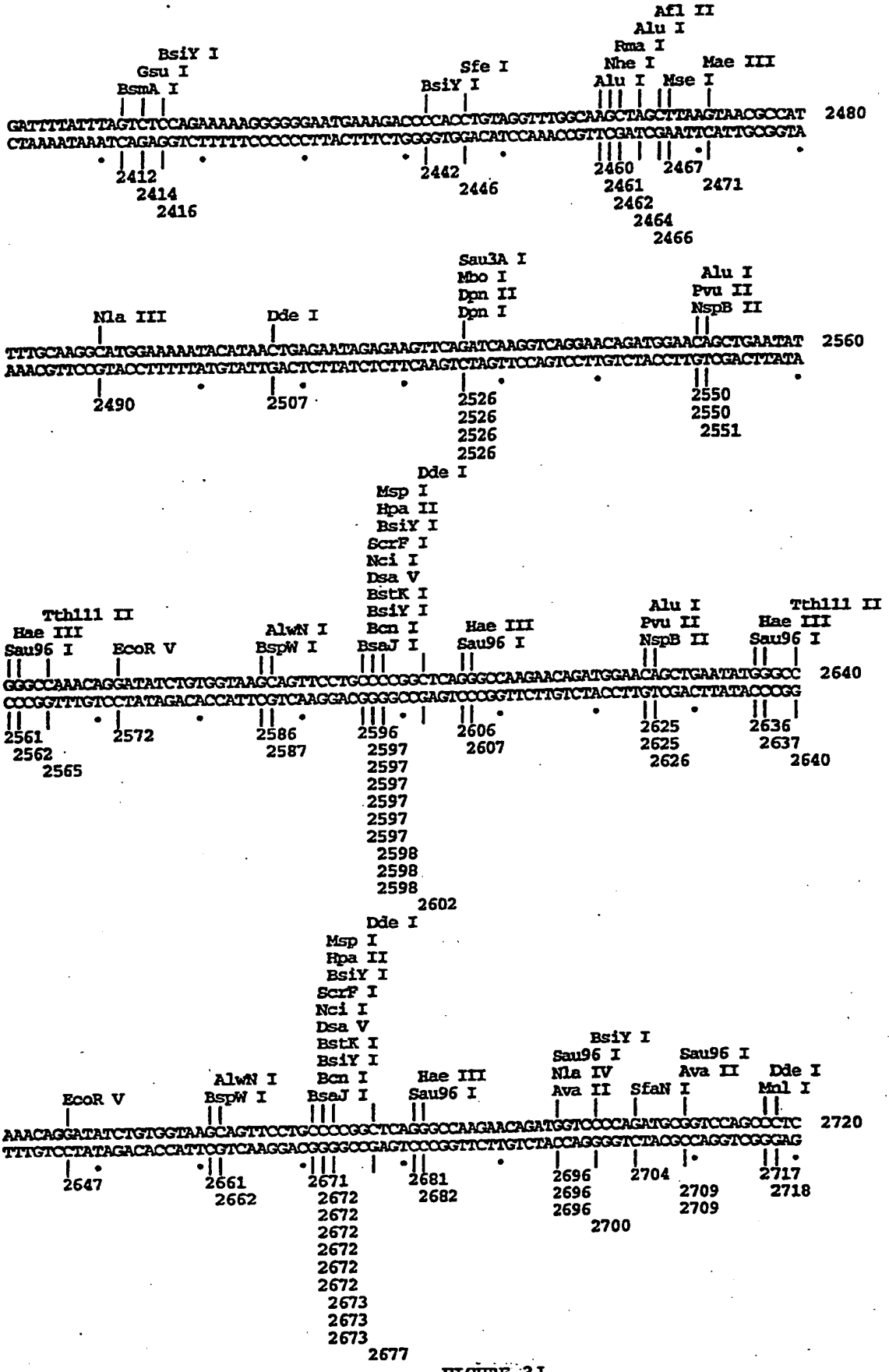

Afl II
Alu I
Rma I
BsiY I
Gsu I
Sfe I
Nhe I
Mae III
BsmA I
BsiY I
Alu I
Mse I
GATTTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCAT 2480
CTAAAATAAATCAGAGGTCTTTTTCCCCCCTTACTTTCTGGGGTGGACATCCAAACGGTTCGATCGAATTCATTGCGGTA
2412
2414
2416
2442
2446
2460
2461
2462
2464
2466
2467
2471
Sau3A I
Mbo I
Dpn II
Dpn I
Alu I
Pvu II
NspB II
Nla III
Dde I
TTTGCAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATAT 2560
AAACGTTCCGTACCTTTTTATGTATTGACTCTTATCTCTTCAAGTCTAGTTCCAGTCCTTGTCTACCTTGTCGACTTATA
2490
2507
2526
2526
2526
2526
2550
2550
2551
Dde I
Msp I
Hpa II
BsiY I
ScrF I
Nci I
Dsa V
BstK I
BsiY I
Tth111 II
Hae III
Sau96 I
EcoR V
AlwN I
BspW I
Bcn I
BsaJ I
Hae III
Sau96 I
Alu I
Pvu II
NspB II
Tth111 II
Hae III
Sau96 I
GGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGAACAGCTGAATATGGGCC 2640
CCCGGTTTGTCCTATAGACACCATTCGTCAAGGACGGGGCCGAGTCCCGGTTCTTGTCTACCTTGTCGACTTATACCCGG
2561
2562
2565
2572
2586
2587
2596
2597
2597
2597
2597
2597
2597
2598
2598
2598
2602
2606
2607
2625
2625
2626
2636
2637
2640
Dde I
Msp I
Hpa II
BsiY I
ScrF I
Nci I
Dsa V
BstK I
BsiY I
BsiY I
Sau96 I
Nla IV
Ava II
Sau96 I
Ava II
Dde I
Mnl I
EcoR V
AlwN I
BspW I
Bcn I
BsaJ I
Hae III
Sau96 I
SfaN I
AAACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTC 2720
TTTGTCCTATAGACACCATTCGTCAAGGACGGGGCCGAGTCCCGGTTCTTGTCTACCAGGGGTCTACGCCAGGTCGGGAG
2647
2661
2662
2671
2672
2672
2672
2672
2672
2672
2673
2673
2673
2677
2681
2682
2696
2696
2696
2700
2704
2709
2709
2717
2718

FIGURE 3J

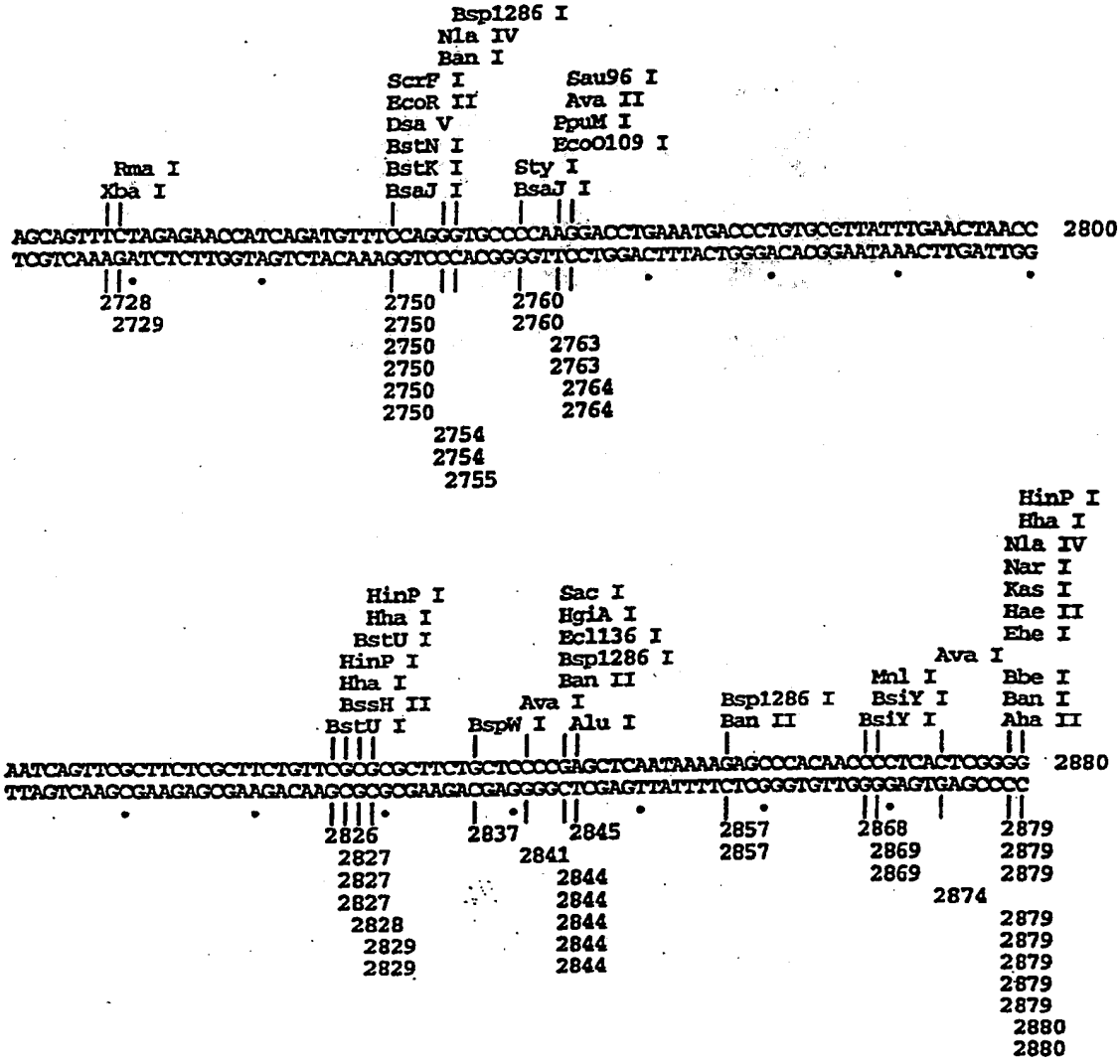

Bsp1286 I
Nla IV
Ban I
ScrF I
EcoR II
Dsa V
BstN I
BstK I
BsaJ I
Sau96 I
Ava II
PpuM I
Eco0109 I
Sty I
BsaJ I
Rma I
Xba I
AGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCCTGTGCCTTATTTGAACTAACC 2800
TCGTCAAAGATCTCTTGGTAGTCTACAAAGGTCCCACGGGGTTCCTGGACTTTACTGGGACACGGAATAAACTTGATTGG
2728
2729
2750
2750
2750
2750
2750
2750
2754
2754
2755
2760
2760
2763
2763
2764
2764
HinP I
Hha I
BstU I
HinP I
Hha I
BssH II
BstU I
Sac I
HgiA I
Ecl136 I
Bsp1286 I
Ban II
Ava I
BspW I
Alu I
Bsp1286 I
Ban II
Mnl I
BsiY I
BsiY I
Ava I
HinP I
Hha I
Nla IV
Nar I
Kas I
Hae III
Ehe I
Bbe I
Ban I
Aha II
AATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAAGAGCCCACAACCCCTCACTCGGGG 2880
TTAGTCAAGCGAAGAGCGAAGACAAGCGCGCGAAGACGAGGGGCTCGAGTTATTTTCTCGGGTGTTGGGGAGTGAGCCCC
2826
2827
2827
2827
2828
2829
2829
2837
2841
2845
2844
2844
2844
2844
2857
2857
2868
2869
2869
2874
2879
2879
2879
2879
2879
2879
2879
2879
2880
2880

FIGURE 3K

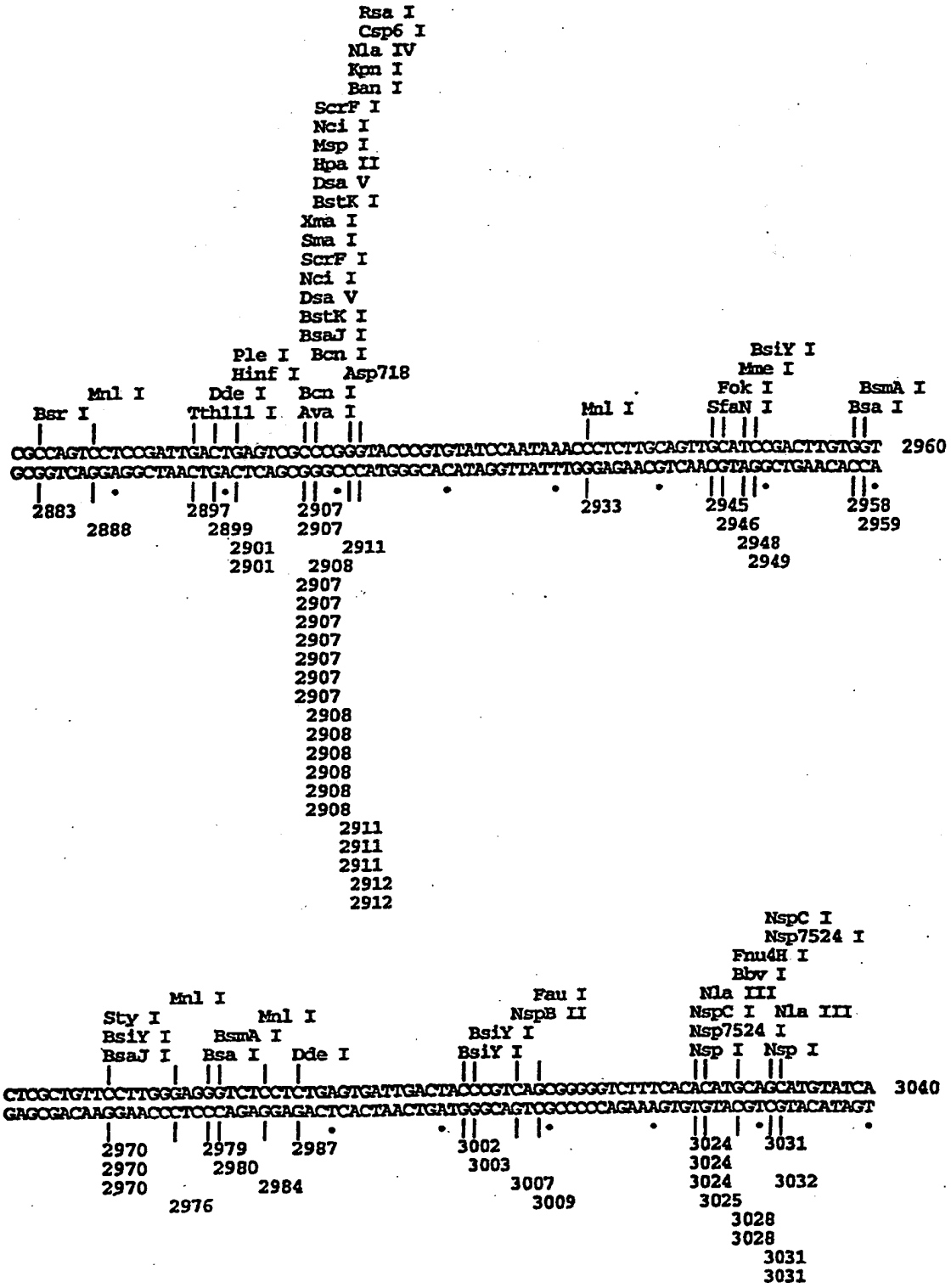

Rsa I
Csp6 I
Nla IV
Kpn I
Ban I
ScrF I
Nci I
Msp I
Hpa II
Dsa V
BstK I
Xma I
Sma I
ScrF I
Nci I
Dsa V
BstK I
BsaJ I
Ple I
Hinf I
Bcn I
Asp718
Mnl I
Dde I
Bcn I
Bsr I
Tth111 I
Ava I
Mnl I
BsiY I
Mme I
Fok I
BsmA I
SfaN I
Bsa I
CGCCAGTGCTCCGATTGACTGAGTCGCCCGGGTACCCGTGTATCCAATAAACCCTCTTGCAGTTGCATCCGACTTGTGGT 2960
GCGGTCAGGAGGCTAACTGACTCAGCGGGCCCATGGGCACATAGGTTATTTGGGAGAACGTCAACGTAGGCTGAACACCA
2883
2888
2897
2899
2901
2901
2907
2907
2911
2908
2907
2907
2907
2907
2907
2907
2908
2908
2908
2908
2908
2908
2911
2911
2911
2912
2912
2933
2945
2946
2948
2949
2958
2959
NspC I
Nsp7524 I
Fnu4H I
Bbv I
Nla III
Mnl I
Sty I
Mnl I
Pau I
NspB II
NspC I
Nla III
BsiY I
BsmA I
BsiY I
Nsp7524 I
BsaJ I
Bsa I
Dde I
BsiY I
Nsp I
Nsp I
CTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGATTGACTACCCGTCAGCGGGGGTCTTTCACACATGCAGCATGTATCA 3040
GAGCGACAAGGAACCCTCCCAGAGGAGACTCACTAACTGATGGGCAGTCGCCCCCAGAAAGTGTGTACGTCGTACATAGT
2970
2970
2970
2976
2979
2980
2984
2987
3002
3003
3007
3009
3024
3024
3024
3025
3028
3028
3031
3031
3031
3032

FIGURE 3L

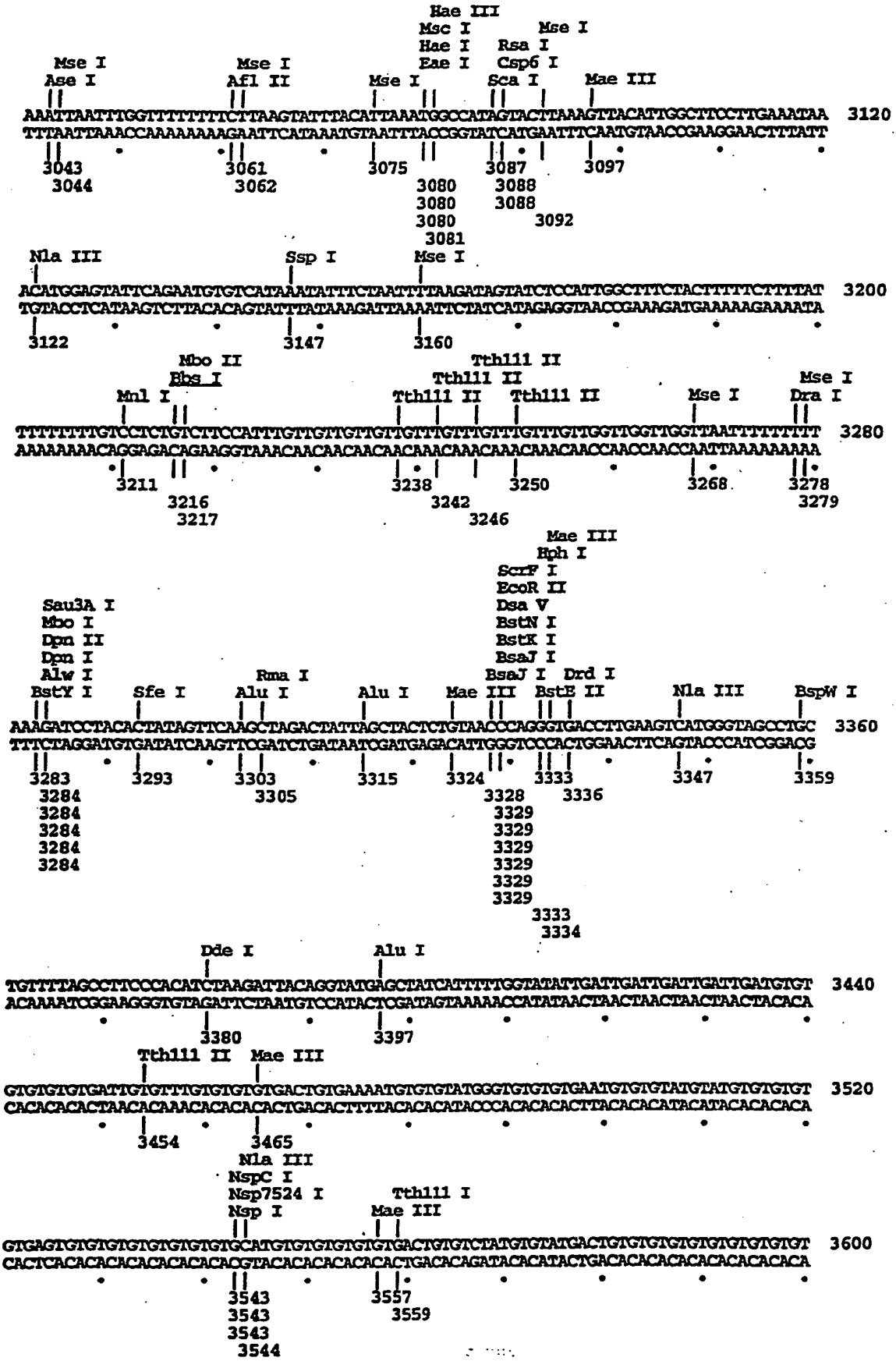

Hae III
Msc I
Hae I
Eae I
Mse I
Rsa I
Csp6 I
Mse I
Ase I
Mse I
Afl II
Mse I
Sca I
Mae III
AAATTAATTTGGTTTTTTTCTTAAGTATTTACATTAAATGGCCATAGTACTTAAAGTTACATTGGCTTCCTTGAAATAA 3120
TTTAATTAAACCAAAAAAAAGAATTCATAAATGTAATTTACCGGTATCATGAATTTCAATGTAACCGAAGGAACTTTATT
3043
3044
3061
3062
3075
3080
3080
3080
3081
3087
3088
3088
3092
3097
Nla III
Ssp I
Mse I
ACATGGAGTATTCAGAATGTGTCATAAATATTTCTAATTTTAAGATAGTATCTCCATTGGCTTTCTACTTTTTCTTTTAT 3200
TGTACCTCATAAGTCTTACACAGTATTTATAAAGATTAAAATTCTATCATAGAGGTAACCGAAAGATGAAAAAGAAAATA
3122
3147
3160
Nbo II
Bbs I
Tth111 II
Tth111 II
Mse I
Mnl I
Tth111 II
Tth111 II
Mse I
Dra I
TTTTTTTGTCCTCTGTCTTCCATTGTTGTTGTTGTTGTTTGTTTGTTGTTGTTGGTTGGTTGGTTAATTTTTTTTT 3280
AAAAAAAACAGGAGACAGAAGGTAAACAACAACAACAACAAACAAACAAACAACCAACCAACCAATTAAAAAAAAA
3211
3216
3217
3238
3242
3246
3250
3268
3278
3279
Mae III
Hph I
ScrF I
EcoR II
Dsa V
BstN I
BstK I
BsaJ I
Sau3A I
Mbo I
Dpn II
Dpn I
Alw I
BstY I
Sfe I
Rma I
Alu I
Alu I
BsaJ I
Drd I
Mae III
BstE II
Nla III
BspW I
AAAGATCCTACACTATAGTTCAAGCTAGACTATTAGCTACTCTGTAACCCAGGGTGACCTTGAAGTCATGGGTAGCCTGC 3360
TTTCTAGGATGTGATATCAAGTTCGATCTGATAATCGATGAGACATTGGGTCCCACTGGAACTTCAGTACCCATCGGACG
3283
3284
3284
3284
3284
3284
3293
3303
3305
3315
3324
3328
3329
3329
3329
3329
3329
3329
3333
3334
3333
3336
3347
3359
Dde I
Alu I
TGTTTTAGCCTTCCCACATCTAAGATTACAGGTATGAGCTATCATTTTTGGTATATTGATTGATTGATTGATTGATGTGT 3440
ACAAAATCGGAAGGGTGTAGATTCTAATGTCCATACTCGATAGTAAAAACCATATAACTAACTAACTAACTAACTACACA
3380
3397
Tth111 II
Mae III
GTGTGTGTGATTGTGTTTGTGTGTGTGACTGTGAAAATGTGTGTATGGGTGTGTGTGAATGTGTGTATGTATGTGTGTGT 3520
CACACACACTAACACAAACACACACACTGACACTTTTACACACATACCCACACACACTTACACACATACATACACACACA
3454
3465
Nla III
NspC I
Nsp7524 I
Nsp I
Tth111 I
Mae III
GTGAGTGTGTGTGTGTGTGTGTGCATGTGTGTGTGTGTGACTGTGTCTATGTGTATGACTGTGTGTGTGTGTGTGTGTGT 3600
CACTCACACACACACACACACACGTACACACACACACACTGACACAGATACACATACTGACACACACACACACACACACA
3543
3543
3543
3544
3557
3559

FIGURE 3M

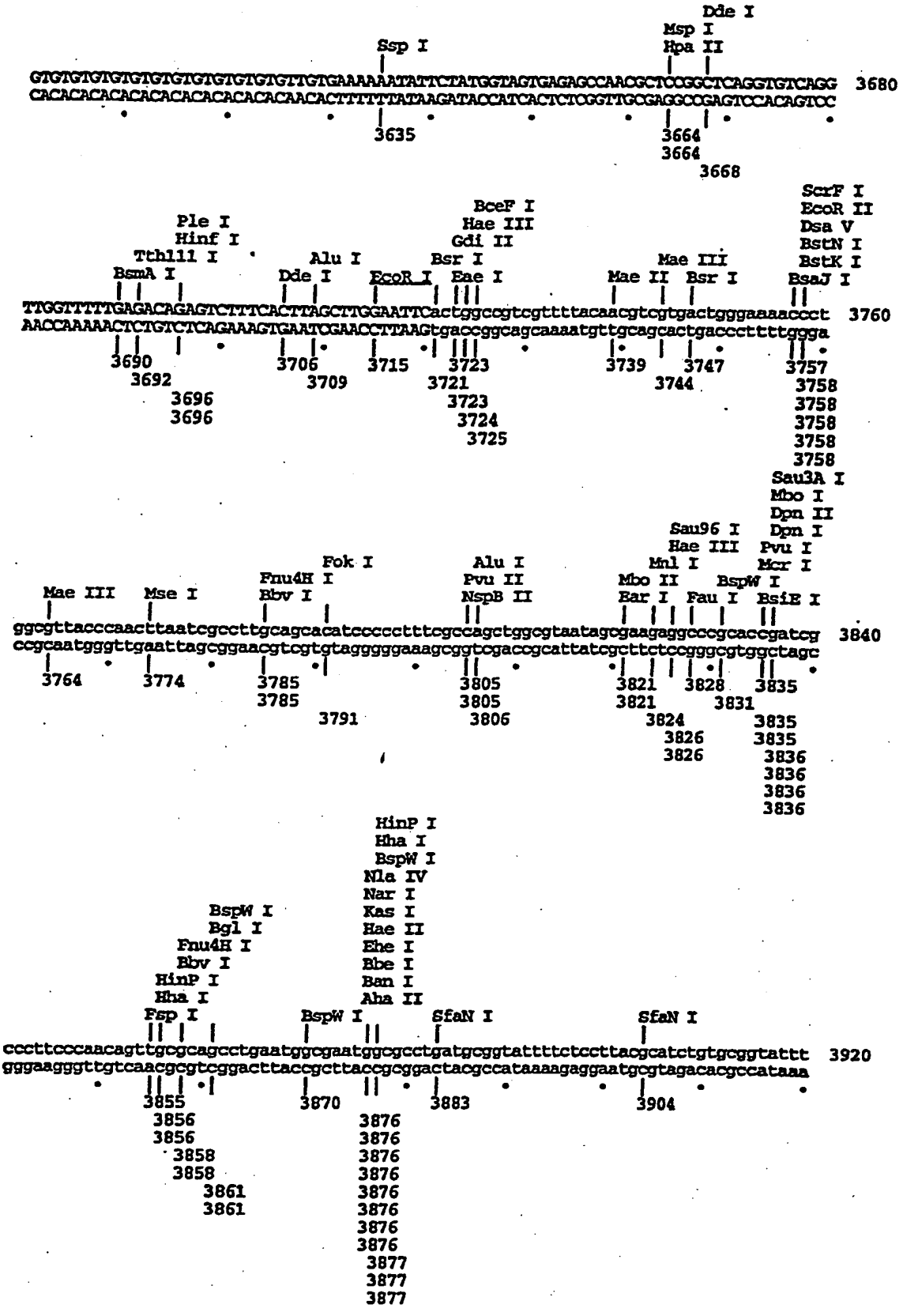

Dde I
Msp I
Hpa II
Ssp I
GTGTGTGTGTGTGTGTGTGTGTGTTGTGAAAAAATATTCTATGGTAGTGAGAGCCAACGCTCCGGCTCAGGTGTCAGG 3680
CACACACACACACACACACACACACACAACACTTTTTTATAAGATACCATCACTCTCGGTTGCGAGGCCGAGTCCACAGTCC
3635
3664
3664
3668
ScrF I
EcoR II
Dsa V
BstN I
BstK I
BsaJ I
BceF I
Hae III
Gdi II
Ple I
Hinf I
Tth111 I
Alu I
Bsr I
Mae III
BsmA I
Dde I
EcoR I
Eae I
Mae II
Bsr I
TTGGTTTTTGAGACAGAGTCTTTCACTTAGCTTGGAATTCactggccgtcgttttacaacgtcgtgactgggaaaaccct 3760
AACCAAAAACTCTGTCTCAGAAAGTGAATCGAACCTTAAGtgaccggcagcaaaatgttgcagcactgacccttttgggа
3690
3692
3696
3696
3706
3709
3715
3723
3721
3723
3724
3725
3739
3744
3747
3757
3758
3758
3758
3758
3758
Sau3A I
Mbo I
Dpn II
Sau96 I
Hae III
Dpn I
Pvu I
Mcr I
Mnl I
Fok I
Alu I
Fnu4H I
Pvu II
Mbo II
BspW I
Mae III
Mse I
Bbv I
NspB II
Bar I
Fau I
BsiE I
ggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcg 3840
ccgcaatgggttgaattagcggaacgtcgtgtaggggaaagcggtcgacccgcattatcgttctccgggcgtggctagc
3764
3774
3785
3785
3791
3805
3805
3806
3821
3821
3824
3826
3826
3828
3831
3835
3835
3835
3836
3836
3836
3836
HinP I
Hha I
BspW I
Nla IV
Nar I
BspW I
Kas I
Bgl I
Hae II
Fnu4H I
Ehe I
Bbv I
Bbe I
HinP I
Ban I
Hha I
Aha II
Fsp I
BspW I
SfaN I
SfaN I
ccctccccaacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtgcggtattt 3920
gggaagggttgtcaacgcgtcggacttaccgcttaccgcggactacgccataaaagaggaatgcgtagacacgccataaa
3855
3856
3856
3858
3858
3861
3861
3870
3876
3876
3876
3876
3876
3876
3876
3876
3876
3877
3877
3877
3883
3904

FIGURE 3N

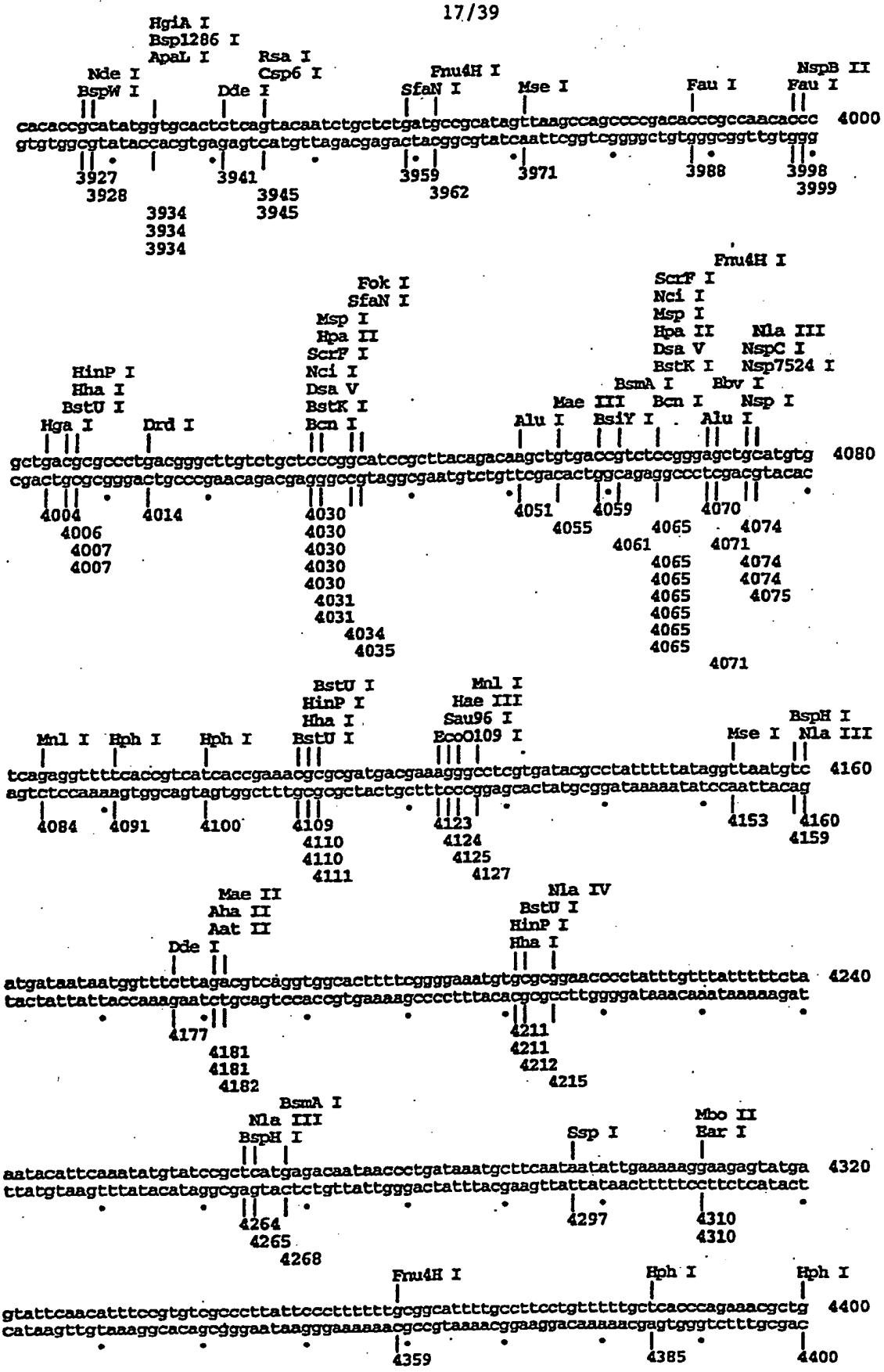

HgiA I
Bsp1286 I
ApaL I
Rsa I
Csp6 I
Nde I
BspW I
Dde I
Fnu4H I
SfaN I
Mse I
Fau I
NspB II
Fau I
cacaccgcatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacaccc 4000
gtgtggcgtataccacgtgagagtcatgttagacgagactacggcgtatcaattcggtcggggctgtgggcggttgtggg
3927
3928
3941
3945
3945
3959
3962
3971
3988
3998
3999
3934
3934
3934
Fnu4H I
ScrF I
Nci I
Msp I
Hpa II
Dsa V
BstK I
Nla III
NspC I
Nsp7524 I
Fok I
SfaN I
Msp I
Hpa II
ScrF I
Nci I
Dsa V
BstK I
Bcn I
HinP I
Hha I
BstU I
Hga I
Drd I
BsmA I
Bbv I
Mae III
Bcn I
Nsp I
Alu I
BsiY I
Alu I
gctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtg 4080
cgactgcgcgggactgcccgaacagacgagggccgtaggcgaatgtctgttcgacactggcagaggccctcgacgtacac
4004
4006
4007
4007
4014
4030
4030
4030
4030
4030
4031
4031
4034
4035
4051
4055
4059
4061
4065
4065
4065
4065
4065
4065
4070
4071
4071
4074
4074
4074
4075
BstU I
HinP I
Hha I
Mnl I
Hae III
Sau96 I
Eco0109 I
BspH I
Nla III
Mnl I
Hph I
Hph I
BstU I
Mse I
tcagaggttttcaccgtcatcaccgaaacgcgcgatgacgaaagggcctcgtgatacgcctatttttataggttaatgtc 4160
agtctccaaaagtggcagtagtggctttgcgcgctactgctttcccggagcactatgcggataaaaatatccaattacag
4084
4091
4100
4109
4110
4110
4111
4123
4124
4125
4127
4153
4160
4159
Mae II
Aha II
Aat II
Nla IV
BstU I
HinP I
Hha I
Dde I
atgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttcta 4240
tactattattaccaaagaatctgcagtccaccgtgaaaagcccctttacacgcgccttggggataaacaaataaaaagat
4177
4181
4181
4182
4211
4211
4212
4215
BsmA I
Nla III
BspH I
Ssp I
Mbo II
Ear I
aatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatga 4320
ttatgtaagtttatacataggcgagtactctgttattgggactatttacgaagttattataactttttcttctcatact
4264
4265
4268
4297
4310
4310
Fnu4H I
Hph I
Hph I
gtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgttttgctcacccagaaacgctg 4400
cataagttgtaaaggcacagcgggaataagggaaaaaacgccgtaaaacggaaggacaaaaacgagtgggtctttgcgac
4359
4385
4400

FIGURE 30

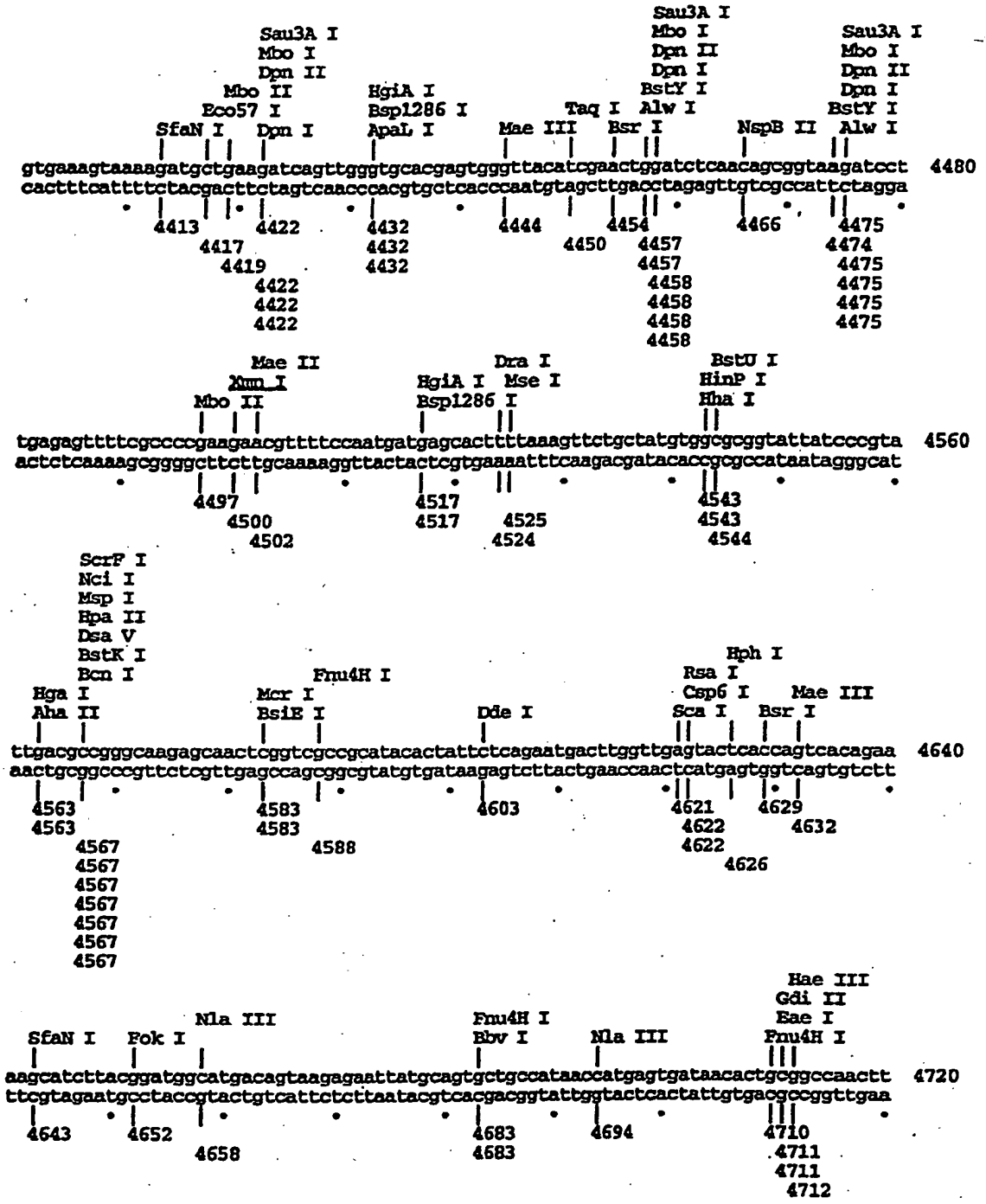

Sau3A I
Mbo I
Dpn II
Mbo II
Eco57 I
SfaN I
Dpn I
HgiA I
Bsp1286 I
ApaL I
Mae III
Taq I
Bsr I
Alw I
BstY I
NspB II
gtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatcct 4480
cactttcatttctacgacttctagtcaacccacgtgctcacccaatgtagcttgacctagagttgtcgccattctagga
4413 4422 4432 4444 4454 4466 4475
4417 4432 4450 4457 4474
4419 4432 4457 4475
4422 4458 4475
4422 4458 4475
4422 4458 4475
4458
Mae II
Xmn I
Mbo II
Dra I
HgiA I
Mse I
Bsp1286 I
BstU I
HinP I
Hha I
tgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgta 4560
actctcaaaagcggggcttcttgcaaaaggttactactcgtgaaaatttcaagacgatacaccgcgccataatagggcat
4497 4517 4543
4500 4517 4525 4543
4502 4524 4544
ScrF I
Nci I
Msp I
Hpa II
Dsa V
BstK I
Bcn I
Hga I
Aha II
Fnu4H I
Mcr I
BsiE I
Dde I
Hph I
Rsa I
Csp6 I
Sca I
Mae III
Bsr I
ttgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaa 4640
aactgcggcccgttctcgttgagccagcggcgtatgtgataagagtcttactgaaccaactcatgagtggtcagtgtctt
4563 4583 4603 4621 4629
4563 4583 4622 4632
4567 4588 4622
4567 4626
4567
4567
4567
4567
4567
Hae III
Gdi II
Eae I
Nla III
SfaN I
Fok I
Fnu4H I
Bbv I
Nla III
Fnu4H I
aagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaactt 4720
ttcgtagaatgcctaccgtactgtcattctcttaatacgtcacgacggtattggtactcactattgtgacgccggttgaa
4643 4652 4683 4694 4710
4658 4683 4711
4711
4712

FIGURE 3P

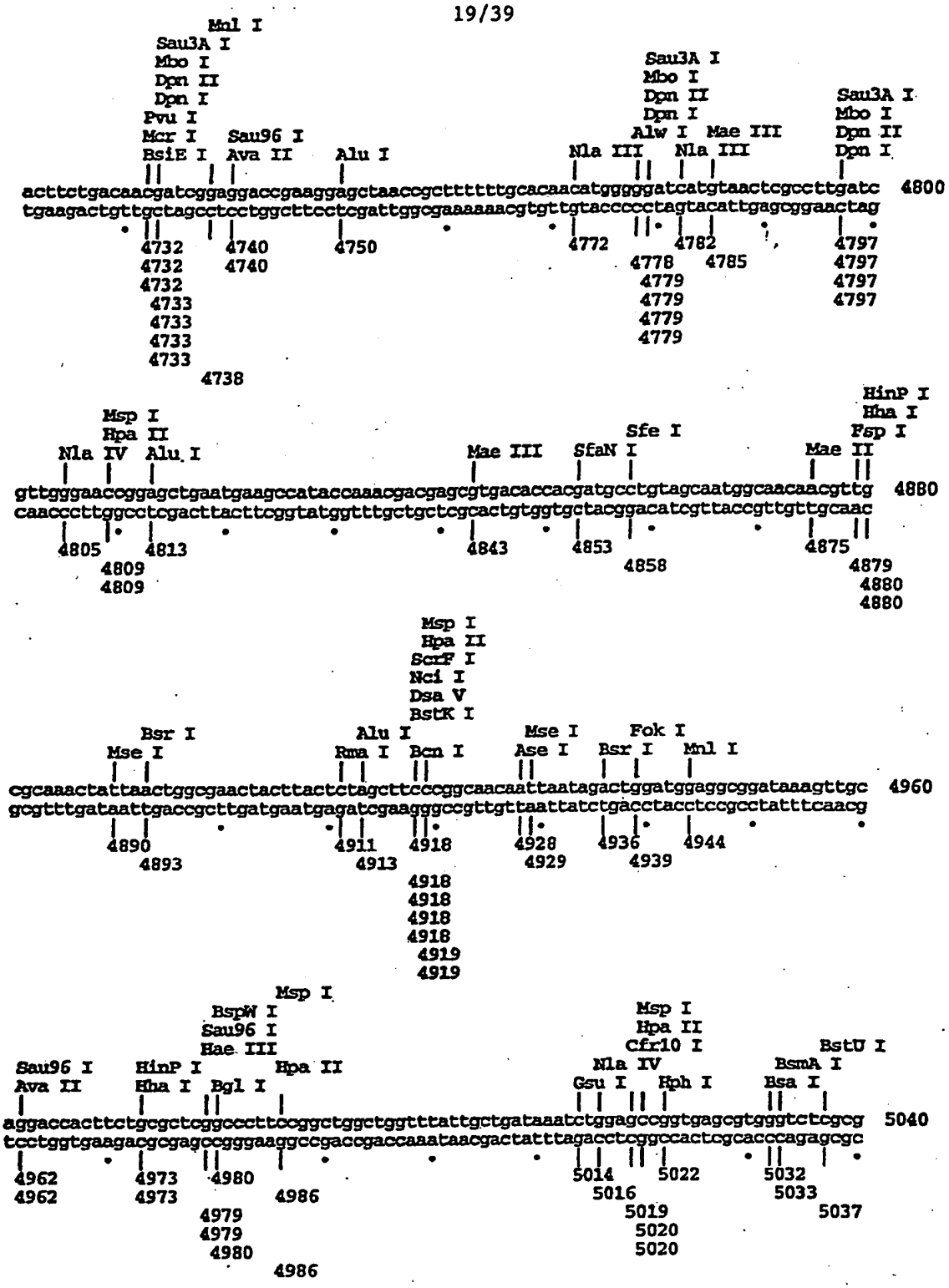

Mnl I
Sau3A I
Mbo I
Dpn II
Dpn I
Pvu I
Mcr I
BsiE I
Sau96 I
Ava II
Alu I
Nla III
Alw I
Mae III
Sau3A I
Mbo I
Dpn II
Dpn I
Sau3A I
Mbo I
Dpn II
Dpn I
acttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatc 4800
tgaagactgttgctagcctcctggcttcctcgattggcgaaaaaacgtgttgtacccccctagtacattgagcggaactag
4732
4732
4732
4733
4733
4733
4733
4738
4740
4740
4750
4772
4778
4779
4779
4779
4779
4782
4785
4797
4797
4797
4797
HinP I
Hha I
Msp I
Hpa II
Nla IV
Alu I
Mae III
SfaN I
Sfe I
Fsp I
Mae II
gttgggaacggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttg 4880
caacccttggcctcgacttacttcggtatggtttgctgctcgcactgtggtgctacggacatcgttaccgttgttgcaac
4805
4809
4809
4813
4843
4853
4858
4875
4879
4880
4880
Msp I
Hpa II
ScrF I
Nci I
Dsa V
BstK I
Bsr I
Mse I
Alu I
Rma I
Bcn I
Mse I
Ase I
Fok I
Bsr I
Mnl I
cgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgc 4960
gcgtttgataattgaccgcttgatgaatgagatcgaagggccgttgttaattatctgacctacctccgcctatttcaacg
4890
4893
4911
4913
4918
4918
4918
4918
4918
4919
4919
4928
4929
4936
4939
4944
Msp I
BspW I
Sau96 I
Hae III
Sau96 I
Ava II
HinP I
Hha I
Bgl I
Hpa II
Msp I
Hpa II
Cfr10 I
Nla IV
Gsu I
Hph I
BstU I
BsmA I
Bsa I
aggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcg 5040
tcctggtgaagacgcgagccgggaaggccgacgaccaaataacgactatttagacctcggccactcgcacccagagcgc
4962
4962
4973
4973
4979
4979
4980
4980
4986
4986
5014
5016
5019
5020
5020
5022
5032
5033
5037

FIGURE 3Q

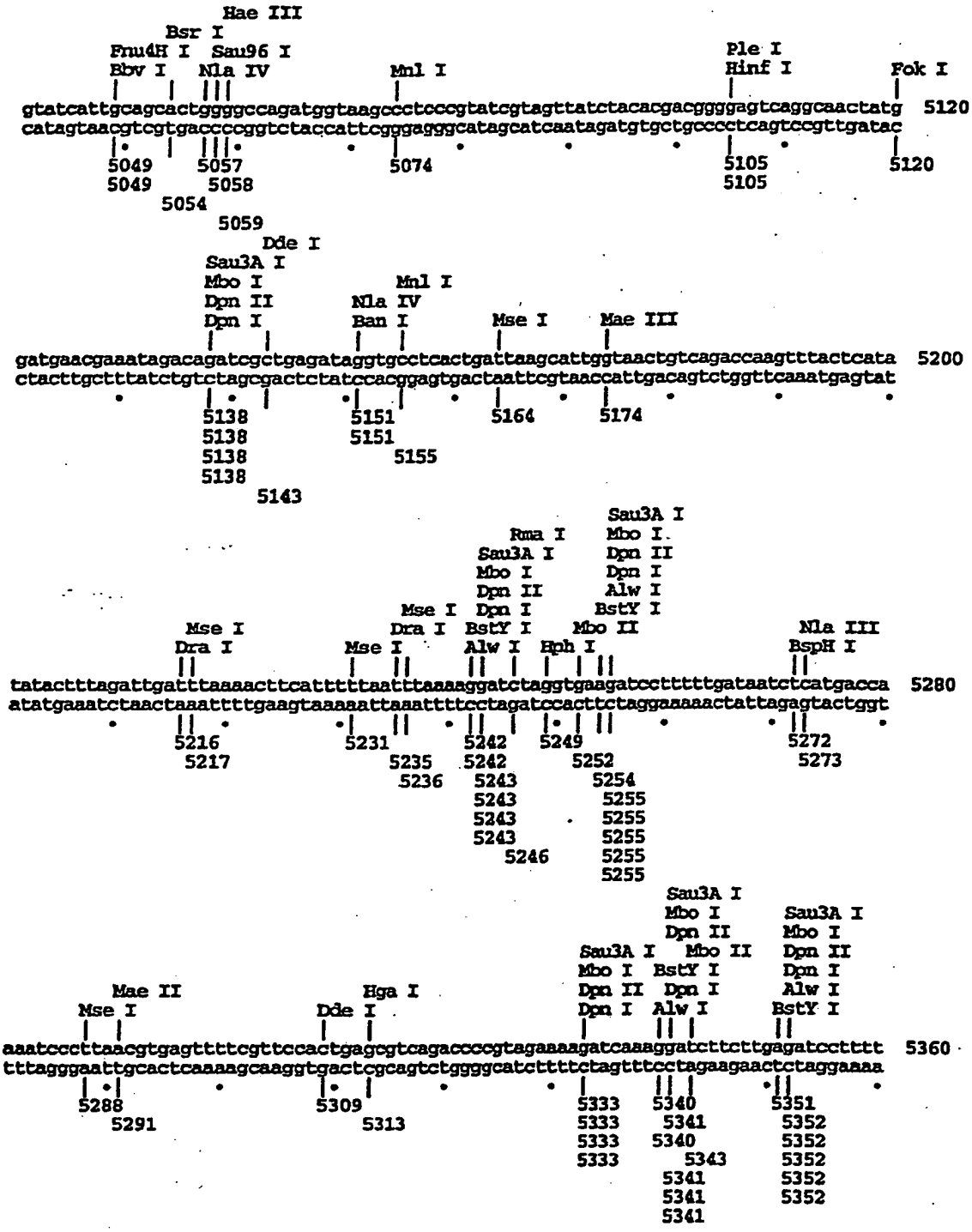

Hae III
Bsr I
Fnu4H I
Sau96 I
Bbv I
Nla IV
Mnl I
Ple I
Hinf I
Fok I
gtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatg 5120
catagtaacgtcgtgaccccggtctaccattcgggaggcatagcatcaatagatgtgctgcccctcagtccgttgatac
5049 5057 5074 5105 5120
5049 5058 5105
5054
5059
Dde I
Sau3A I
Mbo I
Dpn II
Dpn I
Mnl I
Nla IV
Ban I
Mse I
Mae III
gatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcata 5200
ctacttgctttatctgtctagcgactctatccacggagtgactaattcgtaaccattgacagtctggttcaaatgagtat
5138 5151 5164 5174
5138 5151
5138 5155
5138
5143
Sau3A I
Rma I
Mbo I
Sau3A I
Dpn II
Mbo I
Dpn I
Dpn II
Alw I
Mse I
Dpn I
BstY I
Mse I
Dra I
BstY I
Mbo II
Nla III
Dra I
Mse I
Alw I
Bph I
BspH I
tatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttttgataatctcatgacca 5280
atatgaaatctaactaaatttttgaagtaaaaaattaaatttttcctagatccactttctaggaaaaactattagagtactggt
5216 5231 5242 5249 5272
5217 5235 5242 5252 5273
5236 5243 5254
5243 5255
5243 5255
5243 5255
5246 5255
5255
Sau3A I
Mbo I
Sau3A I
Dpn II
Mbo I
Sau3A I
Mbo II
Dpn II
Mbo I
BstY I
Dpn I
Dpn II
Dpn I
Alw I
Dpn I
Alw I
BstY I
Mae II
Mse I
Hga I
Dde I
aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttt 5360
tttagggaattgcactcaaaagcaaggtgactcgcagtctggggcatcttttctagtttcctagaagaactctaggaaaa
5288 5309 5333 5340 5351
5291 5313 5333 5341 5352
5333 5340 5352
5333 5343 5352
5341 5352
5341 5352
5341

FIGURE 3R

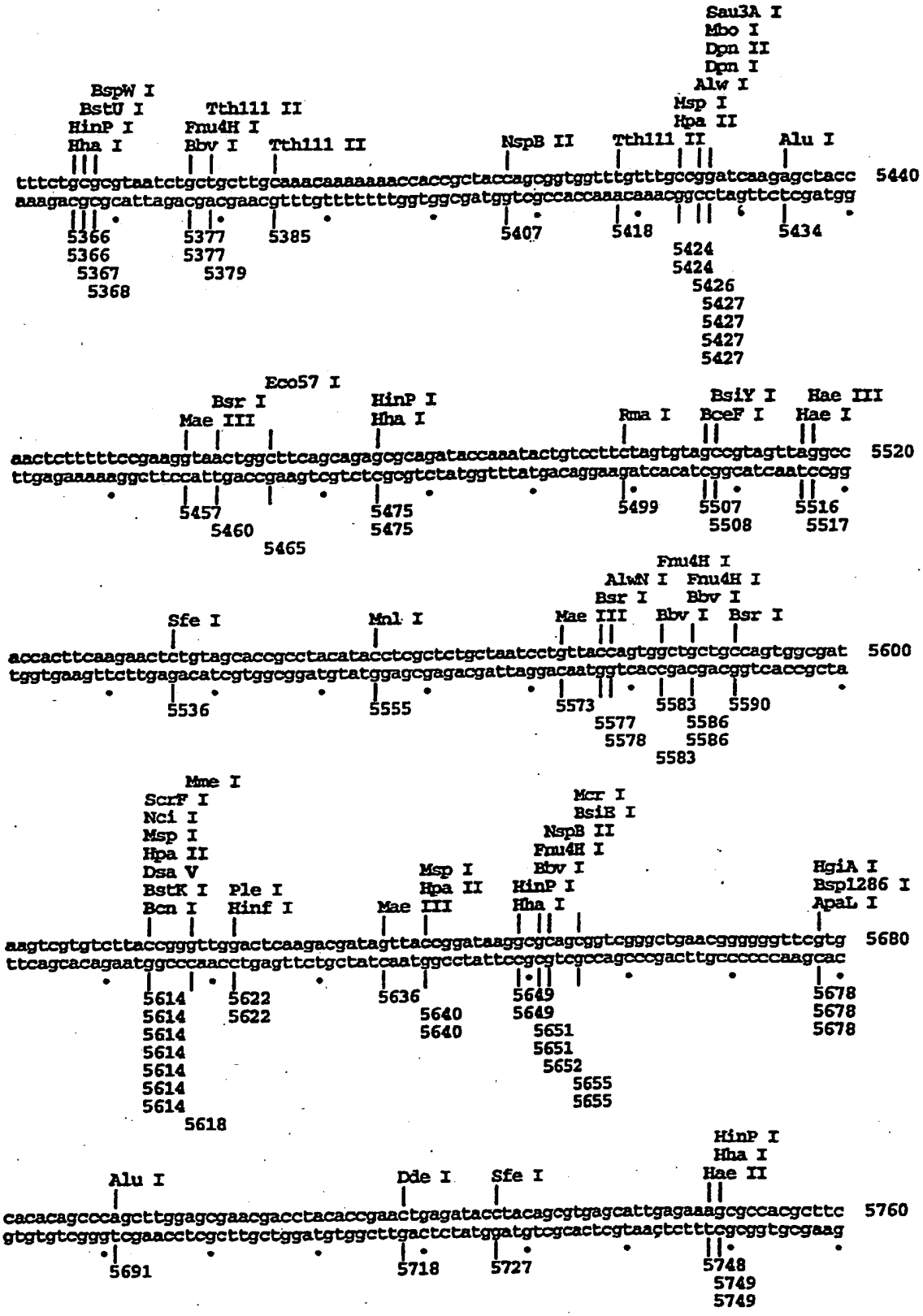

Sau3A I
Mbo I
Dpn II
Dpn I
Alw I
Msp I
Hpa II
BspW I
BstU I
HinP I
Hha I
Tth111 II
Fnu4H I
Bbv I
Tth111 II
NspB II
Tth111 II
Alu I
tttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctacc 5440
aaagacgcgcattagacgacgaacgtttgttttttggtggcgatggtcgccaccaaacaaacggcctagttctcgatgg
5366
5366
5367
5368
5377
5377
5379
5385
5407
5418
5424
5424
5426
5427
5427
5427
5427
5434
Eco57 I
Bsr I
Mae III
HinP I
Hha I
Rma I
BsiY I
BceF I
Hae III
Hae I
aactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggcc 5520
ttgagaaaaaggcttccattgaccgaagtcgtctcgcgtctatggtttatgacaggaagatcacatcggcatcaatccgg
5457
5460
5465
5475
5475
5499
5507
5508
5516
5517
Fnu4H I
AlwN I
Fnu4H I
Bsr I
Bbv I
Sfe I
Mnl I
Mae III
Bbv I
Bsr I
accacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgat 5600
tggtgaagttcttgagacatcgtggcggatgtatggagcgagacgattaggacaatggtcaccgacgacggtcaccgcta
5536
5555
5573
5577
5578
5583
5583
5586
5586
5590
Mme I
ScrF I
Nci I
Msp I
Hpa II
Dsa V
BstK I
Bcn I
Ple I
Hinf I
Msp I
Hpa II
Mae III
Mcr I
BsiE I
NspB II
Fmu4H I
Bbv I
HinP I
Hha I
HgiA I
Bsp1286 I
ApaL I
aagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtg 5680
ttcagcacagaatggcccaacctgagttctgctatcaatggcctattccgcgtcgccagcccgacttgccccccaagcac
5614
5614
5614
5614
5614
5614
5614
5618
5622
5622
5636
5640
5640
5649
5649
5651
5651
5652
5655
5655
5678
5678
5678
HinP I
Hha I
Alu I
Dde I
Sfe I
Hae II
cacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgccacgcttc 5760
gtgtgtcgggtcgaacctcgcttgctggatgtggcttgactctatggatgtcgcactcgtaagtctttcgcggtgcgaag
5691
5718
5727
5748
5749
5749

FIGURE 3S

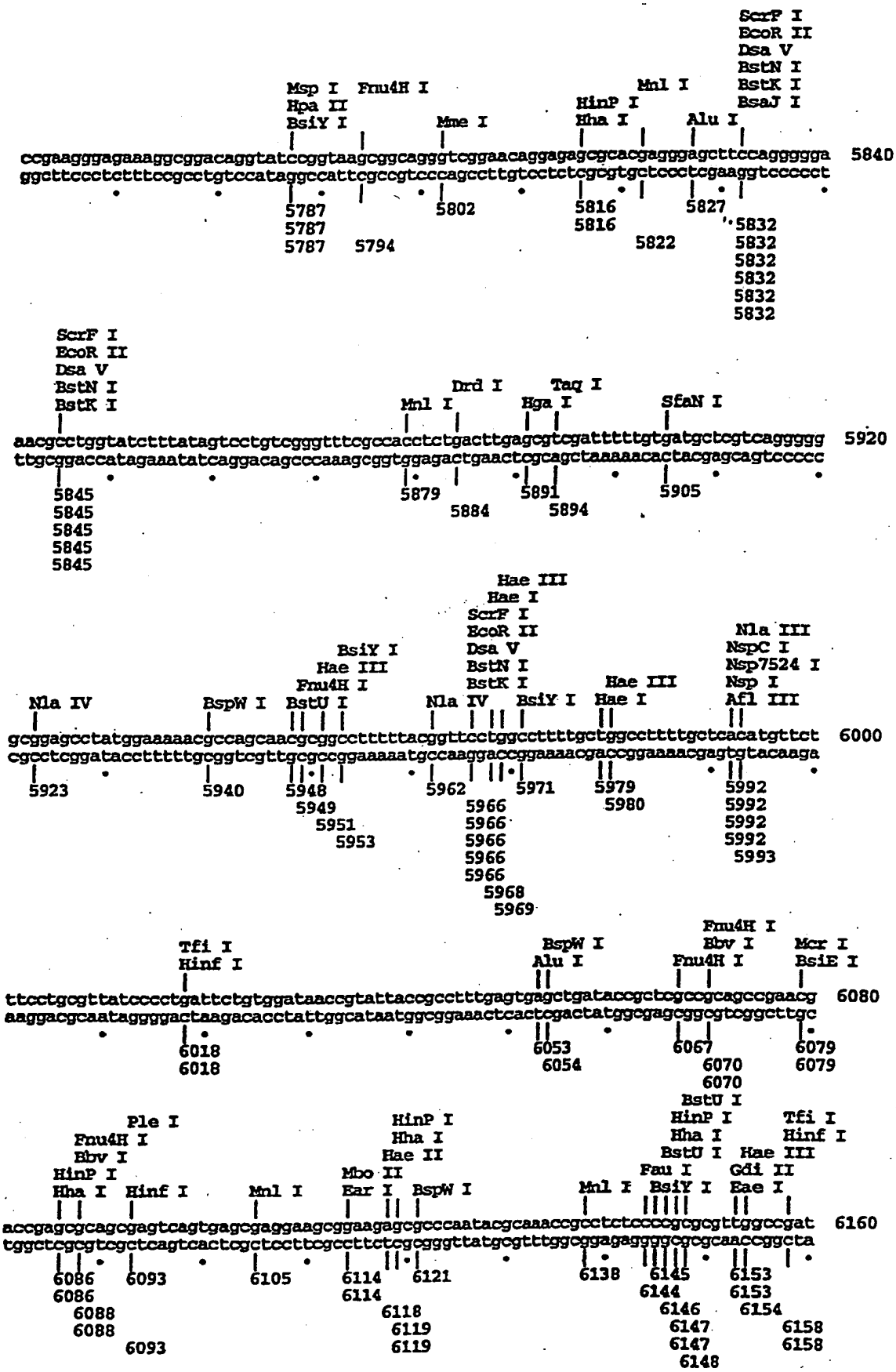

ScrF I
EcoR II
Dsa V
BstN I
BstK I
Msp I
Hpa II
BsiY I
Fnu4H I
Mme I
HinP I
Hha I
Mnl I
Alu I
5840
5787
5802
5816
5827
5822
5794
5832
Drd I
Taq I
Mnl I
Hga I
SfaN I
5920
5845
5879
5891
5905
5884
5894
Hae III
Hae I
Nla III
NspC I
Nsp7524 I
Nsp I
Afl III
Nla IV
BspW I
BstU I
BsiY I
Hae I
6000
5923
5940
5948
5962
5971
5979
5980
5992
5966
5968
5969
Tfi I
Hinf I
Fnu4H I
Bbv I
Mcr I
BsiE I
6080
6018
6053
6054
6067
6070
6079
Ple I
HinP I
Hha I
Hae II
Mbo II
Ear I
Fau I
Gdi II
Eae I
6160
6086
6088
6093
6105
6114
6121
6118
6119
6138
6144
6145
6146
6147
6148
6153
6154
6158

FIGURE 3T

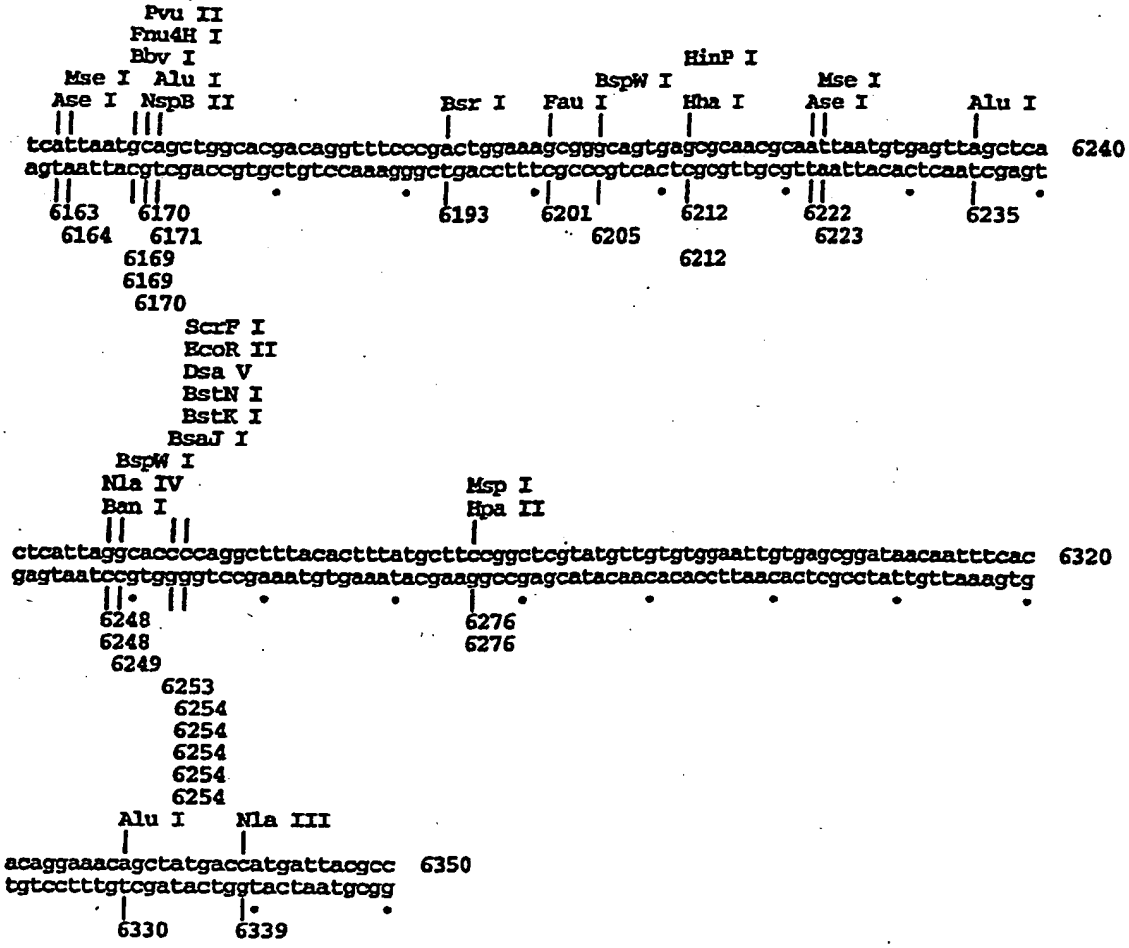

Pvu II
Fnu4H I
Bbv I
HinP I
Mse I Alu I
BspW I
Mse I
Ase I NspB II
Bsr I
Fau I
Hha I
Ase I
Alu I
tcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctca 6240
agtaattacgtcgaccgtgctgtccaaagggctgaccttcgcccgtcactcgcgttgcgttaattacactcaatcgagt
6163 6170 6193 6201 6212 6222 6235
6164 6171 6205 6223
6169 6212
6169
6170
ScrF I
EcoR II
Dsa V
BstN I
BstK I
BsaJ I
BspW I
Nla IV
Msp I
Ban I
Hpa II
ctcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcac 6320
gagtaatccgtggggtccgaaatgtgaaatacgaaggccgagcatacaacacaccttaacactcgcctattgttaaagtg
6248 6276
6248 6276
6249
6253
6254
6254
6254
6254
6254
Alu I Nla III
acaggaaacagctatgaccatgattacgcc 6350
tgtcctttgtcgatactggtactaatgcgg
6330 6339

FIGURE 3U

```
   aagcttgggctgcaggtcgatcgactctagaggatcgatccccaccATGgGTCAATCACG
 1 ---------+---------+---------+---------+---------+---------+ 60
   ttcgaacccgacgtccagctagctgagatctcctagctaggggtggTACcCAGTTAGTGC

                                                  MetGlyGlnSerArg -


   CTACCTCCTCTTTTTGGCCACCCTTGCCCTCCTAAACCACCTCAGTTTGGCCAGGGTCAT
61 ---------+---------+---------+---------+---------+---------+ 120
   GATGGAGGAGAAAAACCGGTGGGAACGGGAGGATTTGGTGGAGTCAAACCGGTCCCAGTA

    TyrLeuLeuPheLeuAlaThrLeuAlaLeuLeuAsnHisLeuSerLeuAlaArgValIle -


   TCCAGTCTCTGGACCTGCCAGGTGTCTTAGCCAGTCCCGAAACCTGCTGAAGACCACAGA
121 ---------+---------+---------+---------+---------+---------+ 180
   AGGTCAGAGACCTGGACGGTCCACAGAATCGGTCAGGGCTTTGGACGACTTCTGGTGTCT

    ProValSerGlyProAlaArgCysLeuSerGlnSerArgAsnLeuLeuLysThrThrAsp -


   TGACATGGTGAAGACGGCCAGAGAAAAACTGAAACATTATTCCTGCACTGCTGAAGACAT
181 ---------+---------+---------+---------+---------+---------+ 240
   ACTGTACCACTTCTGCCGGTCTCTTTTTGACTTTGTAATAAGGACGTGACGACTTCTGTA

    AspMetValLysThrAlaArgGluLysLeuLysHisTyrSerCysThrAlaGluAspIle -
```

FIGURE 4A

```
     CGATCATGAAGACATCACACGGGACCAAACCAGCACATTGAAGACCTGTTTACCACTGGA
241  ---------+---------+---------+---------+---------+---------+ 300
     GCTAGTACTTCTGTAGTGTGCCCTGGTTTGGTCGTGTAACTTCTGGACAAATGGTGACCT

      AspHisGluAspIleThrArgAspGlnThrSerThrLeuLysThrCysLeuProLeuGlu  -


     ACTACACAAGAACGAGAGTTGCCTGGCTACTAGAGAGACTTCTTCCACAACAAGAGGGAG
301  ---------+---------+---------+---------+---------+---------+ 360
     TGATGTGTTCTTGCTCTCAACGGACCGATGATCTCTCTGAAGAAGGTGTTGTTCTCCCTC

      LeuHisLysAsnGluSerCysLeuAlaThrArgGluThrSerSerThrThrArgGlySer  -


     CTGCCTGCCCCCACAGAAGACGTCTTTGATGATGACCCTGTGCCTTGGTAGCATCTATGA
361  ---------+---------+---------+---------+---------+---------+ 420
     GACGGACGGGGGTGTCTTCTGCAGAAACTACTACTGGGACACGGAACCATCGTAGATACT

      CysLeuProProGlnLysThrSerLeuMetMetThrLeuCysLeuGlySerIleTyrGlu  -


     GGACTTGAAGATGTACCAGACAGAGTTCCAGGCCATCAACGCAGCACTTCAGAATCACAA
421  ---------+---------+---------+---------+---------+---------+ 480
     CCTGAACTTCTACATGGTCTGTCTCAAGGTCCGGTAGTTGCGTCGTGAAGTCTTAGTGTT

      AspLeuLysMetTyrGlnThrGluPheGlnAlaIleAsnAlaAlaLeuGlnAsnHisAsn  -


     CCATCAGCAGATCATTCTAGACAAGGGCATGCTGGTGGCCATCGATGAGCTGATGCAGTC
481  ---------+---------+---------+---------+---------+---------+ 540
     GGTAGTCGTCTAGTAAGATCTGTTCCCGTACGACCACCGGTAGCTACTCGACTACGTCAG

      HisGlnGlnIleIleLeuAspLysGlyMetLeuValAlaIleAspGluLeuMetGlnSer -
```

FIGURE 4B

```
    TCTGAATCATAATGGCGAGACTCTGCGCCAGAAACCTCCTGTGGGAGAAGCAGACCCTTA
541 ---------+---------+---------+---------+---------+---------+ 600
    AGACTTAGTATTACCGCTCTGAGACGCGGTCTTTGGAGGACACCCTCTTCGTCTGGGAAT

     LeuAsnHisAsnGlyGluThrLeuArgGlnLysProProValGlyGluAlaAspProTyr  -


    CAGAGTGAAAATGAAGCTCTGCATCCTGCTTCACGCCTTCAGCACCCGCGTCGTGACCAT
601 ---------+---------+---------+---------+---------+---------+ 660
    GTCTCACTTTTACTTCGAGACGTAGGACGAAGTGCGGAAGTCGTGGGCGCAGCACTGGTA

     ArgValLysMetLysLeuCysIleLeuLeuHisAlaPheSerThrArgValValThrIle  -


    CAACAGGGTGATGGGCTATCTGAGCTCCGCCTGAgaattcattgatccactag
661 ---------+---------+---------+---------+---------+---------+ 720
    GTTGTCCCACTACCCGATAGACTCGAGGCGGACTcttaagtaactaggtgatc

     AsnArgValMetGlyTyrLeuSerSerAlaEnd
```

FIGURE 4C

```
    AAGCTTGGGCTGCAGGTCGATCGACTCTAGAGGATCGATCCCCACCATGGGTCCTCAGAA
  1 ---------+---------+---------+---------+---------+---------+ 60
    TTCGAACCCGACGTCCAGCTAGCTGAGATCTCCTAGCTAGGGGTGGTACCCAGGAGTCTT

                                                 Met Gly Pro Gln Lys  -

    GCTAACCATCTCCTGGTTTGCCATCGTTTTGCTGGTGTCTCCACTCATGGCCATGTGGGA
 61 ---------+---------+---------+---------+---------+---------+ 120
    CGATTGGTAGAGGACCAAACGGTAGCAAAACGACCACAGAGGTGAGTACCGGTACACCCT

    LeuThrIleSerTrpPheAlaIleValLeuLeuValSerProLeuMetAlaMetTrpGlu   -

    GCTGGAGAAAGACGTTTATGTTGTAGAGGTGGACTGGACTCCCGATGCCCCTGGAGAAAC
121 ---------+---------+---------+---------+---------+---------+ 180
    CGACCTCTTTCTGCAAATACAACATCTCCACCTGACCTGAGGGCTACGGGGACCTCTTTG

    LeuGluLysAspValTyrValValGluValAspTrpThrProAspAlaProGlyGluThr   -

    AGTGAACCTCACCTGTGACACGCCTGAAGAAGATGACATCACCTGGACCTCAGACCAGAG
181 ---------+---------+---------+---------+---------+---------+ 240
    TCACTTGGAGTGGACACTGTGCGGACTTCTTCTACTGTAGTGGACCTGGAGTCTGGTCTC

    ValAsnLeuThrCysAspThrProGluGluAspAspIleThrTrpThrSerAspGlnArg   -
```

FIGURE 5A

```
    ACATGGAGTCATAGGCTCTGGAAAGACCCTGACCATCACTGTCAAAGAGTTTCTAGATGC
241 ---------+---------+---------+---------+---------+---------+ 300
    TGTACCTCAGTATCCGAGACCTTTCTGGGACTGGTAGTGACAGTTTCTCAAAGATCTACG   -

    HisGlyValIleGlySerGlyLysThrLeuThrIleThrValLysGluPheLeuAspAla   -


    TGGCCAGTACACCTGCCACAAAGGAGGCGAGACTCTGAGCCACTCACATCTGCTGCTCCA
301 ---------+---------+---------+---------+---------+---------+ 360
    ACCGGTCATGTGGACGGTGTTTCCTCCGCTCTGAGACTCGGTGAGTGTAGACGACGAGGT

    GlyGlnTyrThrCysHisLysGlyGlyGluThrLeuSerHisSerHisLeuLeuLeuHis   -


    CAAGAAGGAAAATGGAATTTGGTCCACTGAAATTTTAAAAAATTTCAAAAACAAGACTTT
361 ---------+---------+---------+---------+---------+---------+ 420
    GTTCTTCCTTTTACCTTAAACCAGGTGACTTTAAAATTTTTTAAAGTTTTTGTTCTGAAA

    LysLysGluAsnGlyIleTrpSerThrGluIleLeuLysAsnPheLysAsnLysThrPhe   -


    CCTGAAGTGTGAAGCACCAAATTACTCCGGACGGTTCACGTGCTCATGGCTGGTGCAAAG
421 ---------+---------+---------+---------+---------+---------+ 480
    GGACTTCACACTTCGTGGTTTAATGAGGCCTGCCAAGTGCACGAGTACCGACCACGTTTC

    LeuLysCysGluAlaProAsnTyrSerGlyArgPheThrCysSerTrpLeuValGlnArg   -
```

FIGURE 5B

```
    AAACATGGACTTGAAGTTCAACATCAAGAGCAGTAGCAGTTCCCCTGACTCTCGGGCAGT
481 ---------+---------+---------+---------+---------+---------+ 540
    TTTGTACCTGAACTTCAAGTTGTAGTTCTCGTCATCGTCAAGGGGACTGAGAGCCCGTCA

    AsnMetAspLeuLysPheAsnIleLysSerSerSerSerSerProAspSerArgAlaVal   -


    GACATGTGGAATGGCGTCTCTGTCTGCAGAGAAGGTCACACTGGACCAAAGGGACTATGA
541 ---------+---------+---------+---------+---------+---------+ 600
    CTGTACACCTTACCGCAGAGACAGACGTCTCTTCCAGTGTGACCTGGTTTCCCTGATACT

    ThrCysGlyMetAlaSerLeuSerAlaGluLysValThrLeuAspGlnArgAspTyrGlu   -


    GAAGTATTCAGTGTCCTGCCAGGAGGATGTCACCTGCCCAACTGCCGAGGAGACCCTGCC
601 ---------+---------+---------+---------+---------+---------+ 660
    CTTCATAAGTCACAGGACGGTCCTCCTACAGTGGACGGGTTGACGGCTCCTCTGGGACGG

    LysTyrSerValSerCysGlnGluAspValThrCysProThrAlaGluGluThrLeuPro   -


    CATTGAACTGGCGTTGGAAGCACGGCAGCAGAATAAATATGAGAACTACAGCACCAGCTT
661 ---------+---------+---------+---------+---------+---------+ 720
    GTAACTTGACCGCAACCTTCGTGCCGTCGTCTTATTTATACTCTTGATGTCGTGGTCGAA

    IleGluLeuAlaLeuGluAlaArgGlnGlnAsnLysTyrGluAsnTyrSerThrSerPhe   -


    CTTCATCAGGGACATCATCAAACCAGACCCGCCCAAGAACTTGCAGATGAAGCCTTTGAA
721 ---------+---------+---------+---------+---------+---------+ 780
    GAAGTAGTCCCTGTAGTAGTTTGGTCTGGGCGGGTTCTTGAACGTCTACTTCGGAAACTT

    PheIleArgAspIleIleLysProAspProProLysAsnLeuGlnMetLysProLeuLys   -
```

FIGURE 5C

```
    GAACTCACAGGTGGAGGTCAGCTGGGAGTACCCTGACTCCTGGAGCACTCCCCATTCCTA
781 ---------+---------+---------+---------+---------+---------+ 840
    CTTGAGTGTCCACCTCCAGTCGACCCTCATGGGACTGAGGACCTCGTGAGGGGTAAGGAT

    AsnSerGlnValGluValSerTrpGluTyrProAspSerTrpSerThrProHisSerTyr   -


    CTTCTCCCTCAAGTTCTTTGTTCGAATCCAGCGCAAGAAAGAAAAGATGAAGGAGACAGA
841 ---------+---------+---------+---------+---------+---------+ 900
    GAAGAGGGAGTTCAAGAAACAAGCTTAGGTCGCGTTCTTTCTTTTCTACTTCCTCTGTCT

    PheSerLeuLysPhePheValArgIleGlnArgLysLysGluLysMetLysGluThrGlu   -


    GGAGGGGTGTAACCAGAAAGGTGCGTTCCTCGTAGAGAAGACATCTACCGAAGTCCAATG
901 ---------+---------+---------+---------+---------+---------+ 960
    CCTCCCCACATTGGTCTTTCCACGCAAGGAGCATCTCTTCTGTAGATGGCTTCAGGTTAC

    GluGlyCysAsnGlnLysGlyAlaPheLeuValGluLysThrSerThrGluValGlnCys   -


    CAAAGGCGGGAATGTCTGCGTGCAAGCTCAGGATCGCTATTACAATTCCTCATGCAGCAA
961 ---------+---------+---------+---------+---------+---------+1020
    GTTTCCGCCCTTACAGACGCACGTTCGAGTCCTAGCGATAATGTTAAGGAGTACGTCGTT

    LysGlyGlyAsnValCysValGlnAlaGlnAspArgTyrTyrAsnSerSerCysSerLys   -


    GTGGGCATGTGTTCCCTGCAGGGTCCGATCCTAGGAATTCC
1021---------+---------+---------+---------+- 1061
    CACCCGTACACAAGGGACGTCCCAGGCTAGGATCTTAAGG

    TrpAlaCysValProCysArgValArgSerEnd
```

FIGURE 5D

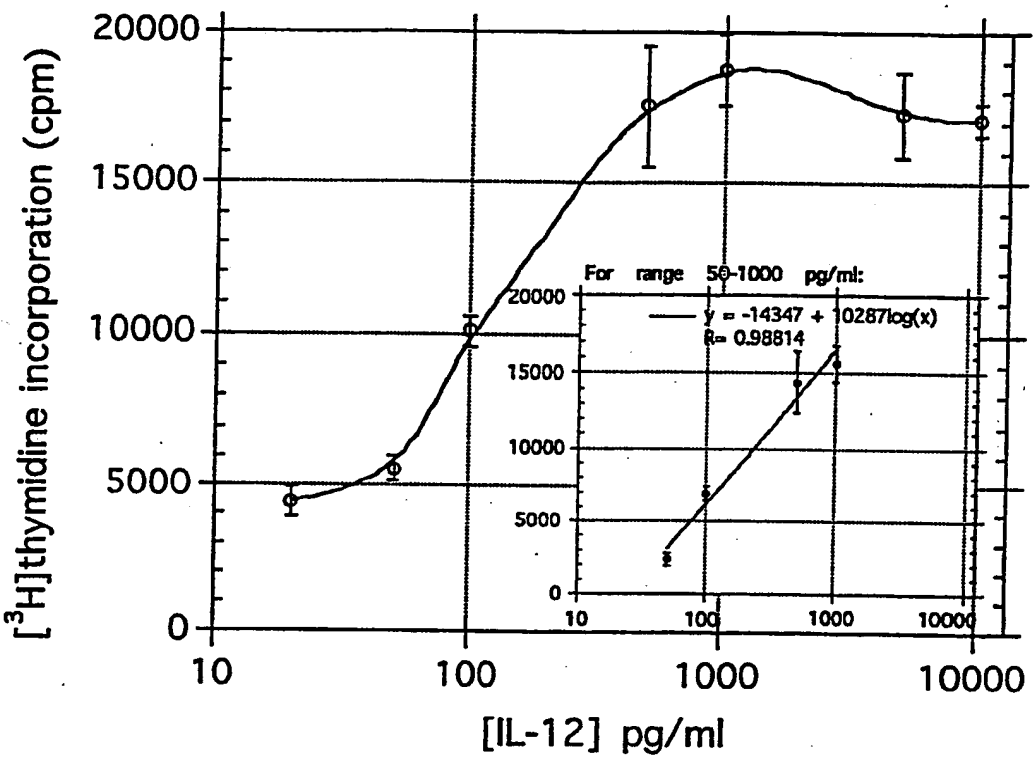

20000
15000
10000
5000
0
[^{3}H]thymidine incorporation (cpm)
10
100
1000
10000
[IL-12] pg/ml
For range 50-1000 pg/ml:
y = -14347 + 10287log(x)
R= 0.98814
20000
15000
10000
5000
0
10
100
1000
10000

FIGURE 6

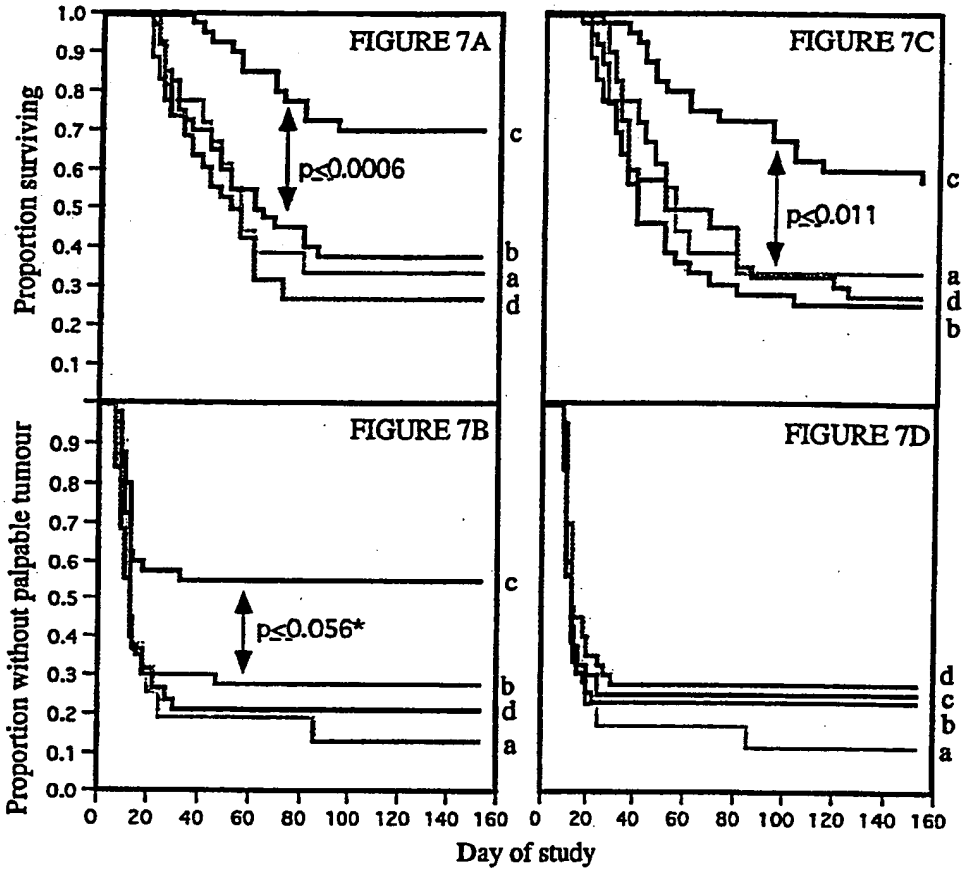


FIGURE 7A
FIGURE 7C
FIGURE 7B
FIGURE 7D
Proportion surviving
Proportion without palpable tumour
p≤0.0006
p≤0.011
p≤0.056*
Day of study
Type of immunotherapy treatment:
a —— no treatment
b —— GM-CSF-secreting cells
c —— IL-12-secreting cells
d —— wild-type cells

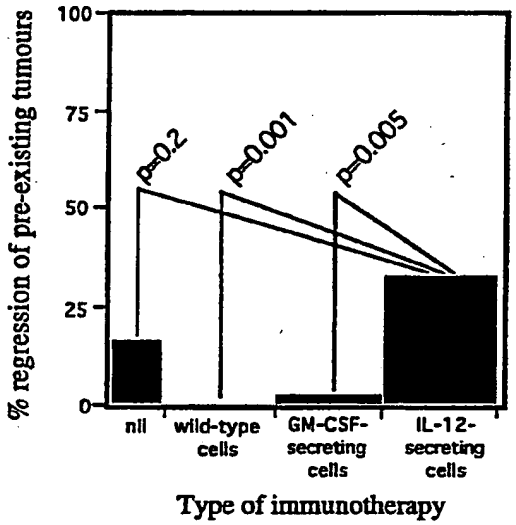

% regression of pre-existing tumours
100
75
50
25
0
p=0.2
p=0.001
p=0.005
nil
wild-type cells
GM-CSF-secreting cells
IL-12-secreting cells
Type of immunotherapy

FIGURE 8

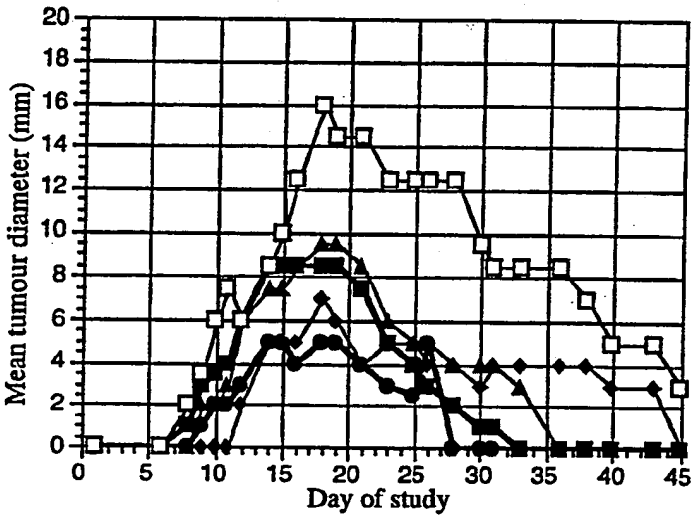
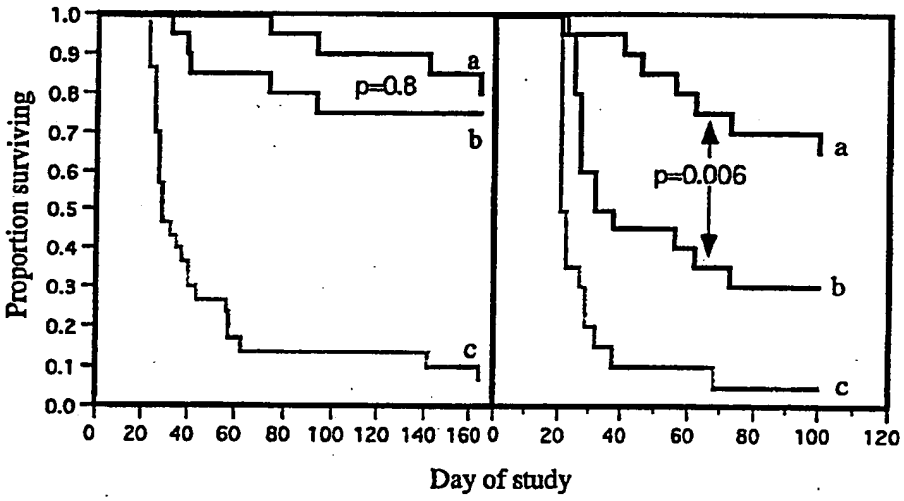

Mean tumour diameter (mm)
20
18
16
14
12
10
8
6
4
2
0
0
5
10
15
20
25
30
35
40
45
Day of study

FIGURE 9

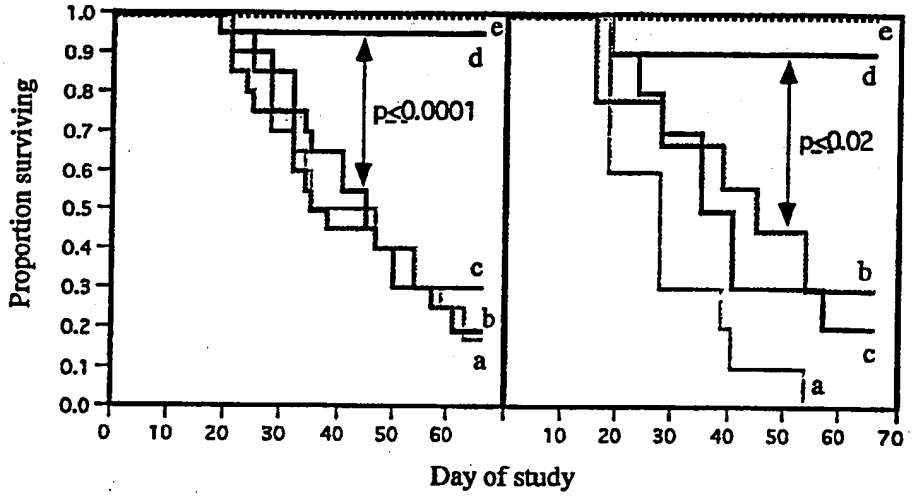
FIGURE 10A FIGURE 10B



Type of Immunotherapy treatment:
a —— no treatment
b —— systemic IL-12
c —— IL-12-secreting cells

FIGURE 11A

FIGURE 11B



Type of immunotherapy treatment:

a —— no treatment
b —— wild-type cells
c —— GM-CSF-secreting cells
d —— IL-12-secreting cells
e ····· IL-12 fusion protein-secreting cells

FIGURE 12A

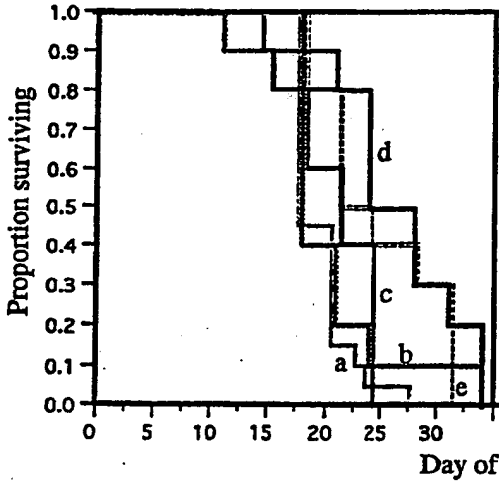


FIGURE 12B

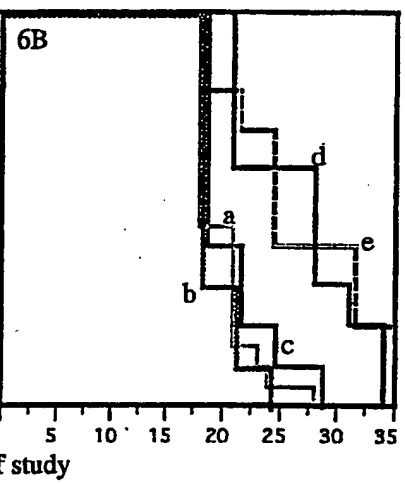
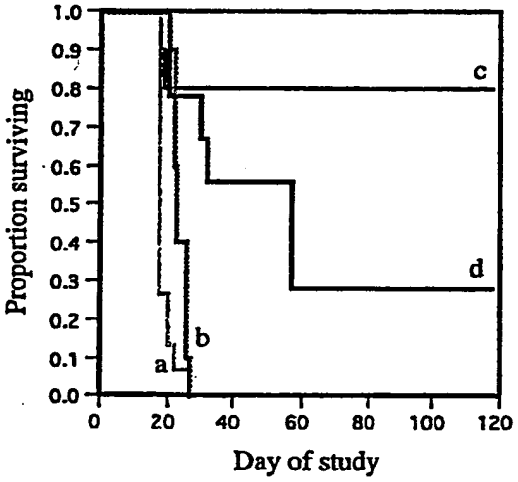


FIGURE 12C



Type of immunotherapy treatment:
- a —— no treatment
- b —— wild-type cells
- c —— GM-CSF-secreting cells
- d —— IL-12-secreting cells
- e ····· IL-12 fusion protein-secreting cells

FIGURE 13A

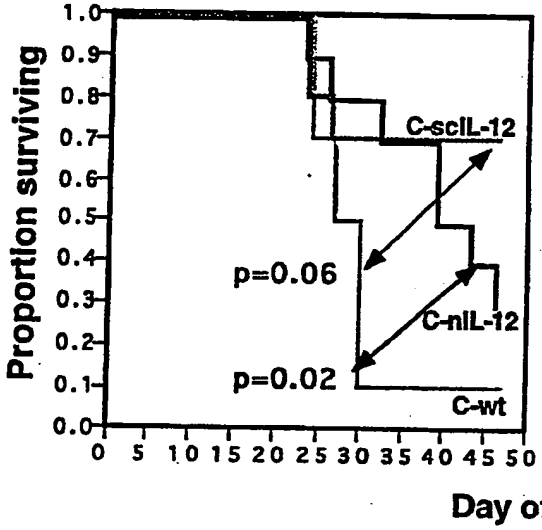


FIGURE 13B

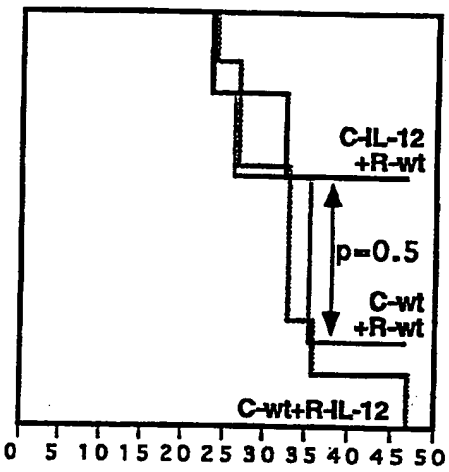


FIGURE 14A

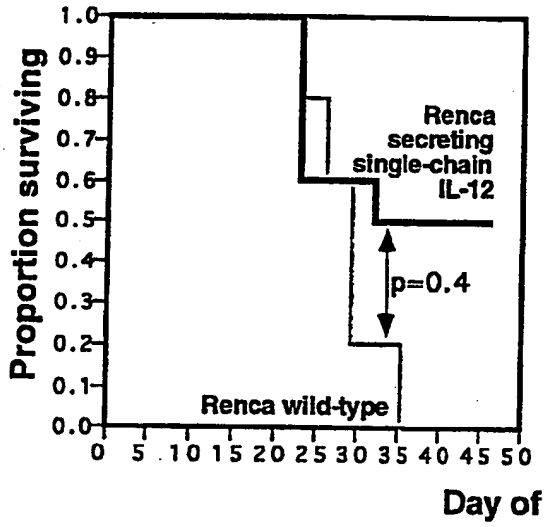
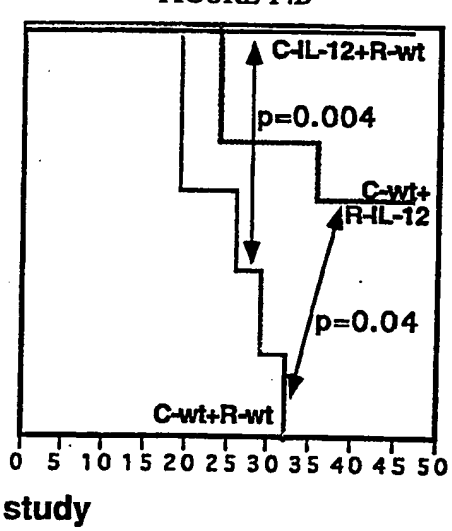


FIGURE 14B



INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 96/01787

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/62 C07K14/54 C07K19/00 A61K48/00 C12N15/24
C12N15/83 A61K38/20

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EUROPEAN JOURNAL OF IMMUNOLOGY , vol. 25, no. 1, January 1995, pages 137-146, XP000574005 A. MARTINOTTI ET AL: "CD4 T cells inhibit in vivo the CD8-mediated immune response against murine colon carcinoma cells transduced with Interleukin-12 genes" see the whole document	23,24, 27,35
Y		1-4, 7-14, 17-20

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 June 1996	09.07.96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Le Cornec, N

INTERNATIONAL SEARCH REPORT

In tional Application No
PCT/US 96/01787

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 85, August 1988, WASHINGTON US, pages 5879-5883, XP002006553 J.S. HUSTON ET AL: "Protein engineering of Antibody binding sites : recovery of specific activity in anti-digoxin single-chain Fv analogue produced in Escherichia coli" cited in the application	5,6
Y	see the whole document	1-4, 7-14, 17-20

Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 88, May 1991, WASHINGTON US, pages 4143-4147, XP002006554 U. GUBLER ET AL: "Coexpression of two distinct genes is required to generate secreted bioactive cytotoxic lymphocyte maturation factor" see the whole document	1-4, 7-14, 17-20

X	EP,A,0 614 982 (ENIRICERCHE S.P.A.) 14 September 1994	5,6
Y	see abstract; claims; figure 3	1-4, 7-14, 17-20

X	WO,A,94 13806 (THE DOW CHEMICAL COMPANY) 23 June 1994	5,6
Y	see abstract; figures 9,10	1-4, 7-14, 17-20
	see page 2	

A	EP,A,0 433 827 (F.HOFFMANN-LA ROCHE AG) 26 June 1991 see claims 16,30	23

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/01787

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Please see Further Information sheet enclosed.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Ir stional Application No
PCT/US 96/01787

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-614982	14-09-94	JP-A-	7075581	20-03-95
WO-A-9413806	23-06-94	AU-B-	5747794	04-07-94
		CA-A-	2117477	23-06-94
		EP-A-	0628078	14-12-94
		JP-T-	7503622	20-04-95
EP-A-433827	26-06-91	AU-B-	5471294	09-06-94
		AU-B-	6834990	27-06-91
		CA-A-	2032653	23-06-91
		JP-A-	5294999	09-11-93